

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: July 1, 2005, 18:05:05 ; Search time 34.3148 Seconds
(without alignments)
1386.327 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_123
Perfect score: 701
Sequence: 1 MSDNQGNNQYQYQSQNG.....FNYNNLQYQAGFPQSQG 123

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	701	100.0	685	4	AAB30792 Amino aci
2	701	100.0	685	6	ABR53107 Protein s
3	701	100.0	685	7	ADK63022 Disease t
4	682	97.3	271	4	AAB30800 A modifie
5	450	64.2	215	4	AAB30799 A modifie
6	228	32.5	715	4	AAB30820 Amino aci
7	223	31.8	721	5	ABP73437 Candida a
8	167	23.8	592	6	ABO53050 Human put
9	167	23.8	592	8	ADI26209 Human pro
10	166	23.7	405	4	AAB30810 Amino aci
11	166	23.7	405	4	AAB30821 Amino aci
12	166	23.7	741	4	AAB30819 Amino aci
13	163.5	23.3	589	8	ADI26207 Human pro
14	163	23.3	1145	4	AAE09770 Drosophil
15	163	23.3	1145	7	ADB38846 Drosophil
16	160	22.8	407	5	AAE94879 ORF incor
17	148.5	21.2	207	5	ABP63118 FLO11 gen
18	148	21.1	417	3	AGS30013 Arabidops
19	148	21.1	419	3	AGS30012 Arabidops
20	148	21.1	439	3	AGS30011 Arabidops
21	143	20.4	518	8	ADI26113 Human pro
22	141	20.1	391	8	ADN03904 Antipsori
23	141	20.1	392	4	ABG10550 Novel hum
24	141	20.1	404	2	AAE90675 Human chr
25	141	20.1	536	4	AAE78356 Human pro

26 139.5 19.9 492 7 ADJ69335 Human hea
27 139 19.8 632 8 ADN73047 Thale cre
28 138.5 19.8 525 5 ABG95084 Human tra
29 138.5 19.8 525 8 ADR14649 Human NF-
30 136.5 19.5 461 2 AAR75506 Nicotiana
31 136 19.4 462 5 ABG95083 Human tra
32 136 19.4 462 8 ADE77159 Human pro
33 136 19.4 526 4 AAM78355 Human pro
34 136 19.4 526 5 ABG95081 Human tra
35 136 19.4 526 8 ADI26117 Human pro
36 136 19.4 526 8 ABM81732 Humour-as
37 136 19.4 526 8 ADS88302 Human pro
38 135.5 19.3 345 4 ABG10551 Novel hum
39 135 19.3 128 4 AAB30822 Amino aci
40 135 19.3 128 8 ADM47991 Polypepti
41 134 19.1 341 2 AAR06522 Tapetum-s
42 134 19.1 341 2 AAR09298 Male flow
43 134 19.1 341 2 AAR09307 Male flow
44 134 19.1 341 2 AAW10628 Anther-ep
45 132 18.8 437 2 AAR75507 Nicotiana

ALIGNMENTS

RESULT 1

AAB30792
ID AAB30792 standard; protein; 685 AA.

XX
AC AAB30792;

XX
DT 02-APR-2001 (first entry)

XX
DE Amino acid sequence of a yeast Sup35 protein.

XX
KW Sup35; phenotype; SCHAG; self-coalesce; higher-order aggregate;
KW amyloidogenic domain; aggregation; fibril; phenotypic alteration;
KW gene therapy; disease resistance; plant pigmentation; prion disease.

XX
OS Saccharomyces cerevisiae.

XX
PN WO200075324-A2.

XX
PD 14-DEC-2000.

XX
PF 09-JUN-2000; 2000WO-US015876.

XX
PR 09-JUN-1999; 99US-0138833P.

XX
XX (ARCH-) ARCH DEV CORP.

XX
PI Lindquist S, Li L, Ma J, Liu J, Sondheimer N, Scheibel T;

XX
DR WPI: 2001-061723/07.

XX
DR N-PSDB; AAC86679.

XX
PT New nucleic acid encoding chimeric proteins with self-assembly
PT properties, useful e.g. for diagnosis and treatment of prion diseases,
PT also related aggregates, fibrils and polymers.

XX
PS Claim 11; Page 125-127; 188pp; English.

XX
CC The present sequence represents a yeast Sup35 protein. The protein
CC possesses the prion-like capacity to undergo a self-perpetuating
CC conformational alteration that changes the functional state of Sup35 in a
CC manner that creates a heritable change in phenotype. It is used to
CC construct chimeric polypeptides of the invention, which comprise at least
CC one SCHAG (self-coalesce into higher-order aggregates) amino acid
CC sequence fused in frame with a polypeptide of interest (which is other
CC than a marker protein, a glutathione-S-transferase or a staphylococcal
CC nuclear protein). The specification also describes chimeric polypeptides
CC that comprises an amyloidogenic domain that causes aggregation into
CC fibrils. The chimeric polypeptides are used to prepare polymers with

BEST AVAILABLE COPY

CC components of the complex is useful for the manufacture of a medicament
CC for the treatment or prevention of a disease or disorder. This sequence
CC corresponds to a protein of the invention. (Note: the sequence data for
CC this patent did not form part of the printed specification but was
CC obtained from the EPO in electronic format).

XX SQ Sequence 685 AA;

Query Match 100.0%; Score 701; DB 7; Length 685;
Best Local Similarity 100.0%; Pred. No. 4.4e-63;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSDSNQGNQNYQQYSSQNGNQGNRRYQGYQAYNAQAQAGGYQNYQYSGYQCGGY 60
Db |||||||
Qy 1 MSDSNQGNQNYQQYSSQNGNQGNRRYQGYQAYNAQAQAGGYQNYQYSGYQCGGY 60
Db |||||||

Qy 61 QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNYKNFYNNNLQYQAGFPQ 120
Db |||||

Qy 61 QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNYKNFYNNNLQYQAGFPQ 120
Db |||||

Qy 121 SQG 123
Db |||||

Qy 121 SQG 123
Db |||||

RESULT 4
AAB30800
ID AAB30800 standard; protein; 271 AA.

XX AC AAB30800;

XX DT 02-APR-2001 (first entry)

XX DE A modified N region of yeast Sup35 protein.

XX KW Sup35; phenotype; SCHAG; self-coalesce; higher-order aggregate;
KW amyloidogenic domain; aggregation; fibril; phenotypic alteration;
KW gene therapy; disease resistance; plant pigmentation; prion disease.

XX OS Synthetic.
OS Saccharomyces cerevisiae.

XX PN WO200075324-A2.
XX 14-DEC-2000.

XX PF 09-JUN-2000; 2000WO-US015876.

XX PR 09-JUN-1999; 99US-0138833P.

XX PA (ARCH-) ARCH DEV CORP.

XX PI Lindquist S, Li L, Ma J, Liu J, Sondheimer N, Scheibel T;
XX WPI; 2001-061723/07.
XX DR N-PSDB; AAC86685.

XX PT New nucleic acid encoding chimeric proteins with self-assembly
XX properties, useful e.g. for diagnosis and treatment of prion diseases,
XX also related aggregates, fibrils and polymers.

XX PS Claim 11; Page 135-136; 188pp; English.

XX CC The present sequence represents a modified N region of Sup35 protein, in
CC the second oligopeptide repeat has been expanded twice, creating a total
CC of seven repeats. Sup35 possesses the prion-like capacity to undergo a
CC self-perpetuating conformational alteration that changes the functional
CC state of Sup35 in a manner that creates a heritable change in phenotype.
CC It is used to construct chimeric polypeptides of the invention, which
CC comprise at least one SCHAG (self-coalesces into higher-order aggregates)
CC amino acid sequence fused in frame with a polypeptide of interest (which
CC is other than a marker protein, a glutathione-S-transferase or a
CC staphylococcal nuclear protein). The specification also describes

CC chimeric polypeptides that comprises an amyloidogenic domain that causes
CC aggregation into fibrils. The chimeric polypeptides are used to prepare
CC polymers with multiple reactivities, e.g. derivatised with enzymes, or
CC specific binding partners, and useful e.g. for performing multi-step
CC chemical reactions. They can be used to create an inducible, or stable
CC phenotypic alteration in a cell, e.g. for gene therapy, protein
CC production, imparting disease resistance to plants, altering plant
CC pigmentation and for diagnosis and treatment of prion diseases

XX SQ Sequence 271 AA;

Query Match 97.3%; Score 682; DB 4; Length 271;
Best Local Similarity 87.2%; Pred. No. 1.2e-61;
Matches 123; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

Qy 1 MSDSNQGNQNYQQYSSQNGNQGNRRYQGYQAYNAQAQAGGYQNYQYSGYQCGGY 56
Db |||||||

Qy 1 MSDSNQGNQNYQQYSSQNGNQGNRRYQGYQAYNAQAQAGGYQNYQYSGYQCGGY 60
Db |||||||

Qy 57 -----QGYQQYNPDAGYQQQYNPQGGYQQQFNPQGGRGNYK 102
Db |||||||

Qy 61 QQYNPQGGYQQYNPQGGYQQYNPDAGYQQQYNPQGGYQQQFNPQGGRGNYK 120
Db |||||||

Qy 103 NFYNNNLQYQAGFPQSQG 123
Db |||||||

Qy 121 NFYNNNLQYQAGFPQSQG 141
Db |||||||

RESULT 5
AAB30799
ID AAB30799 standard; protein; 215 AA.

XX AC AAB30799;

XX DT 02-APR-2001 (first entry)

XX DE A modified N region of yeast Sup35 protein.

XX KW Sup35; phenotype; SCHAG; self-coalesce; higher-order aggregate;
KW amyloidogenic domain; aggregation; fibril; phenotypic alteration;
KW gene therapy; disease resistance; plant pigmentation; prion disease.

XX OS Synthetic.
OS Saccharomyces cerevisiae.

XX PN WO200075324-A2.
XX 14-DEC-2000.

XX PF 09-JUN-2000; 2000WO-US015876.

XX PR 09-JUN-1999; 99US-0138833P.

XX PA (ARCH-) ARCH DEV CORP.

XX PI Lindquist S, Li L, Ma J, Liu J, Sondheimer N, Scheibel T;
XX WPI; 2001-061723/07.
XX DR N-PSDB; AAC86684.

XX PT New nucleic acid encoding chimeric proteins with self-assembly
XX properties, useful e.g. for diagnosis and treatment of prion diseases,
XX also related aggregates, fibrils and polymers.

XX PS Example 3; Page 133-134; 188pp; English.

XX CC The present sequence represents a modified N region of a yeast Sup35
CC protein, in which four of the five oligopeptide repeats have been
CC deleted. Sup35 possesses the prion-like capacity to undergo a self-
CC perpetuating conformational alteration that changes the functional state
CC of Sup35 in a manner that creates a heritable change in phenotype. It is
CC used to construct chimeric polypeptides of the invention, which comprise
CC at least one SCHAG (self-coalesces into higher-order aggregates) amino

subarachnoid haemorrhage; viral hepatitis; AIDS.
Homo sapiens.
WO2003104277-A2.
18-DEC-2003.
05-JUN-2003; 2003WO-JP007123.
05-JUN-2002; 2002JP-00164257.
06-JUN-2002; 2002US-0385912P.
26-DEC-2002; 2002JP-00377326.
27-DEC-2002; 2002US-0436467P.
15-MAY-2003; 2003JP-00137505.
16-MAY-2003; 2003US-0470836P.
(ASAH) ASAH KASEI KK.
Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
WPI; 2004-122214/12.
N-PSDB; ADI26206.
New signal transducer and activator of transcription 6 activation
promoting purified protein, for diagnosing and treating disease
associated with activation/inhibition of transcription factor e.g.
diabetes and cancer.
Claim 1; SEQ ID NO 172; 1368pp; English.
The invention relates to a purified protein promoting signal transducer
and activator of transcription 6 activation (STAT6). The protein is
useful for the producing an antibody, which involves administering the
protein or its epitope-bearing fragments to a non-human animal as an
antigen. The nucleic acid is useful for diagnosing a disease or
susceptibility to a disease related to expression or activity of the
protein. A transformant expressing the protein is useful for screening
compounds which inhibit or promote STAT6 activation. A transformant
expressing the protein is useful for producing a pharmaceutical
composition. Compositions, antibodies and antisense molecules are useful
for the treating a disease associated with STAT6 activation such as
allergic diseases, inflammation, autoimmune diseases, diabetes,
hyperlipidaemia, infectious disease and cancers. Compositions are useful
for treating disease associated with STAT6 activation and/or prevention
of the hyperactive diseases. Compositions are also useful in rheumatoid
arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
viral hepatitis and AIDS. The protein has efficient promoting STAT6
activity. The protein or nucleic acid is effectively useful for screening
compounds for treating and preventing disease associated with excessive
activation or inhibition of STAT6. The present sequence represents the
amino acid sequence of a human protein which promotes STAT6 activation.
Sequence 589 AA;
Query Match 23.3%; Score 163.5; DB 8; Length 589;
Best Local Similarity 30.8%; Pred. No. 4.7e-08;
Matches 49; Conservative 13; Mismatches 42; Indels 55; Gaps 8;
QY 1 MSBDS-----NQGNNQNYQQYSQNGNQGNRRYGGYQAYNAQAQPA-GGYQNYQYSGY 55
Db 1 MSBDSGYSQSGGQSGQSYTYGNGFSQGYG-----QASQSYSGYQYTTDSSYQNYSSY 56
QY 56 QQ-----GGY-----QQYN-----PDAGYQQQYNNPQ 76
Db 57 GQSYSSYSGYENKQSSYSQQFYNNQGGQQNNMESSGSGGRAPSYDQFYGQHSYDQ 116
QY 77 GGYQQYNNPQGYQQQNFQGGRGYKYNFNYYNNLQYQA 115
Db 117 SGYDQH--GQSYDEQSN-----YDQHDYSQNNQSYHS 148

RESULT 14
AAE09770
ID AAE09770 standard; protein; 1145 AA.
XX AC AAE09770;
XX DT 29-NOV-2001 (first entry)
XX DE Drosophila melanogaster argonaute 2 protein.
XX KW Fruit fly; gene expression; dicer; argonaute 2; RNA interference; RNAi;
XX KW attenuation; gene function.
XX OS Drosophila melanogaster.
XX FH Key Location/Qualifiers
FT Region 437..446
FT /note= "Peptide identified by microsequencing"
FT Region 456..462
FT /note= "Peptide identified by microsequencing"
FT Region 485..488
FT /note= "Peptide identified by microsequencing"
FT Region 548..558
FT /note= "Peptide identified by microsequencing"
FT Region 593..605
FT /note= "Peptide identified by microsequencing"
FT Region 740..746
FT /note= "Peptide identified by microsequencing"
FT Region 1112..1119
FT /note= "Peptide identified by microsequencing"
XX WO200168836-A2.
XX PD 20-SEP-2001.
XX PF 16-MAR-2001; 2001WO-US008435.
XX PR 16-MAR-2000; 2000US-0189739P.
XX PR 24-OCT-2000; 2000US-0243097P.
XX (GENE-) GENETICA INC.
XX PA (COLD-) COLD SPRING HARBOR LAB.
XX PI Beach D, Bernstein E, Caudy A, Hammond S, Hannon G;
XX WPI; 2001-565793/63.
XX Attenuating gene expression in a cell using gene-targeted double stranded
XX RNA.
XX Claim 6; Fig 24; 135pp; English.
XX The invention relates to methods for attenuating gene expression in a
XX cell using gene targeted double stranded RNA. The double stranded RNA
XX comprises a nucleotide sequence that hybridises under physiologic
XX conditions of the cell to the DNA sequence of the target gene to be
XX inhibited. The method utilises a cell in which dicer or argonaute
XX activities are recombinantly expressed or otherwise ectopically
XX activated. Activated RNA interference (RNAi) enzymes such as dicer and
XX argonaute are specifically and potentially used for inactivating a cloned
XX gene and proves to be a powerful tool for investigating gene function.
XX Methods are used for attenuating target gene expression in non-embryonic
XX cells suspended in culture or in animals. The present sequence is
XX Drosophila melanogaster argonaute 2 protein related to the invention
XX Sequence 1145 AA;
Query Match 23.3%; Score 163; DB 4; Length 1145;
Best Local Similarity 34.2%; Pred. No. 1.2e-07;
Matches 52; Conservative 10; Mismatches 56; Indels 34; Gaps 7;
QY 6 QGNNQQ-----NYQQY---SQNGNQGNRRYGGYQAYNAQAQPAAGGYQNYQGY--- 52

Db 119 QGGHQGRGQGDGGYQQRPPGQGGHQQGRGQEGGYQQRPPGQGGHQQGRGQEGG 178
Qy 53 -----SGYQGGYQ--YNPAGYQQ--YNPGGYQ--YNPGGYQ--YNPGGYQ--Q 91
Db 179 YQRPSPGQGGHQQGRGQEGGYQQRPPGQGGHQQGRGQEGGYQQRPPGQGGHQQ 238
Qy 92 FNPQGGNGNYKNFNNNLQGYQAGFQPSOG 123
Db 239 QGRGQEGGYQQRPPGQGGHQQGRGQEGG 270

RESULT 15
ADB38846
ID ADB38846 standard; protein; 1145 AA.
XX AC ADB38846;
DT 04-DEC-2003 (first entry)
XX DE Drosophila argonaute 2 protein related to gene expression inhibitors.
KW attenuating expression; double stranded RNA; dsRNA; RNA interference;
KW post-transcriptional gene silencing; quelling; short inhibitory RNA;
KW siRNA; short hairpin RNA; shRNA; gene expression; gene-targeted dsRNA;
KW adult stem cell; Major Histocompatibility Complex; MHC;
KW cellular medicament; Dicer; RNAi enzyme.
XX OS Drosophila sp.
XX WO2003062394-A2.
XX PD 31-JUL-2003.
XX PF 22-JAN-2003; 2003WO-US001963.
XX PR 22-JAN-2002; 2002US-00055797.
XX PA (GENE-) GENETICA INC.
XX PA (COLD-) COLD SPRING HARBOR LAB.
XX PI Beach DH, Bernstein E, Caudy A, Hammond S, Hannon GJ;
XX PI Paddison PJ, Conklin D;
XX WPI; 2003-636734/60.

XX Attenuating expression of a target gene in host cells comprises
XX introducing double stranded RNA into the host cells in an amount that
XX attenuates expression of the target gene.

XX Disclosure: Fig 24; 150pp; English.
XX This invention relates to a novel method of attenuating expression of a
XX target gene in host cells which comprises introducing double stranded RNA
XX (dsRNA) into the host cells in an amount sufficient to attenuate
XX expression of the target gene, where the dsRNA comprises a nucleotide
XX sequence that hybridises under stringent conditions to an untranslated or
XX intronic sequence of the target gene. The method may be known as RNA
XX interference, post-transcriptional gene silencing or quelling and makes
XX use of short inhibitory RNA (siRNA) or short hairpin RNA (shRNA)
XX sequences. The method may be useful for attenuating gene expression in a
XX cell using gene-targeted dsRNA. In a preferred embodiment of the
XX invention, the method is used on undifferentiated adult stem cells to
XX increase or decrease Major Histocompatibility Complex (MHC) expression.
XX The stem cell is useful in manufacturing a cellular medicament for
XX transplantation to a patient. The composition may be used in
XX manufacturing a medicament for attenuating expression of one or more
XX genes in vivo. The present sequence is that of the Drosophila Argonaute 2
XX protein which is related to the invention and a homologue of the human
XX Dicer protein. Dicer, an RNAi enzyme, is required for shRNA mediated gene
XX silencing.

XX Sequence 1145 AA;
SQ

Query Match 23.3%; Score 163; DB 7; Length 1145;
Best Local Similarity 34.2%; Pred. No. 1.2e-07;
Matches 52; Conservative 10; Mismatches 56; Indels 34; Gaps 7;
Qy 6 QGNNQQ-----NYQY---SQNGNQGGNNRYQGYQAYNAQAQAPAGGYQNTYQY--- 52
Db 119 QGGHQGRGQGDGGYQQRPPGQGGHQQGRGQEGGYQQRPPGQGGHQQGRGQEGG 178
Qy 53 -----SGYQGGYQ--YNPAGYQQ--YNPGGYQ--YNPGGYQ--YNPGGYQ--Q 91
Db 179 YQRPSPGQGGHQQGRGQEGGYQQRPPGQGGHQQGRGQEGGYQQRPPGQGGHQQGRGQEGG 238
Qy 92 FNPQGGNGNYKNFNNNLQGYQAGFQPSOG 123
Db 239 QGRGQEGGYQQRPPGQGGHQQGRGQEGG 270

Search completed: July 1, 2005, 18:17:01
Job time : 37.3148 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (C) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 18:06:15 ; Search time 7.18756 Seconds
(without alignments)
1646.548 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_123
Perfect score: 701
Sequence: 1 MSDSNQGNQYQYQSQNG.....FVYNNLQYQAGFPQSQG 123

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	701	100.0	685	1 EFBS22	suppressor 2 prote
2	166	23.7	405	2 S19355	hypothetical prote
3	166	23.7	741	2 S12921	suppressor 2 prote
4	163.5	23.3	589	2 S71954	RNA/esDNA-binding
5	149	21.3	456	2 F88493	protein F57B9.9 [i
6	148	21.1	419	2 T04886	DAG protein homolo
7	147.5	21.0	409	2 T20847	hypothetical prote
8	147.5	21.0	940	2 D89723	protein F39D8.1b [
9	147.5	21.0	945	2 T21998	hypothetical prote
10	147	21.0	462	1 LUD07	annexin VII - slim
11	144	20.5	232	2 T21735	hypothetical prote
12	143	20.4	372	2 I64223	bifunctional endo-
13	142	20.3	676	2 S41022	hypothetical prote
14	141	20.1	404	2 S46269	synovial sarcoma t
15	141	20.1	800	2 I51653	dsRNA-binding prot
16	140.5	20.0	236	2 T18240	hypothetical trans
17	139	19.8	632	2 T02627	hypothetical prote
18	138	19.7	1111	2 T29070	hypothetical prote
19	137.5	19.6	284	2 T23158	hypothetical prote
20	137.5	19.6	765	2 T15447	hypothetical prote
21	137	19.5	455	2 C29349	hypothetical prote
22	136.5	19.5	461	2 T10265	arabingalactan-pr
23	136	19.4	462	4 S33798	FUS/CHOP mutant fu
24	136	19.4	526	1 S33799	RNA-binding protei
25	135	19.3	128	2 S45871	probable membrane
26	135	19.3	528	2 G02127	fus-like protein -
27	135	19.3	867	2 T27136	hypothetical prote
28	135	19.3	871	2 T27135	hypothetical prote
29	133	19.0	3498	2 T22330	hypothetical prote

30	132	18.8	882	2 H70199	translation initia
31	131	18.7	1161	2 T18400	glutamate-cysteine
32	128	18.3	618	2 T19407	hypothetical prote
33	128	18.3	635	2 T19403	hypothetical prote
34	126.5	18.0	738	2 S37876	glutamine-rich pro
35	124.5	17.8	382	2 T293339	hypothetical prote
36	124.5	17.8	1196	2 S65245	translation elonga
37	123	17.5	346	2 S44874	ZC21.3 protein - C
38	121.5	17.3	353	1 S56750	single stranded D
39	121	17.3	597	2 T51889	related to clathri
40	121	17.3	954	1 S20907	endo-1,4-beta-xyla
41	120.5	17.2	989	2 T02568	hypothetical prote
42	119.5	17.0	301	2 JW0079	heterogeneous nucl
43	119.5	17.0	557	2 T19685	hypothetical prote
44	119.5	17.0	1139	2 T32275	hypothetical prote
45	118.5	16.9	129	2 T06256	dormancy-associate

ALIGNMENTS

RESULT 1

EFBS22
suppressor 2 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: G1-to-S transition protein; protein YD9395.05; protein YDR172w
C:Species: Saccharomyces cerevisiae
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: S00733; S49768; S00488; A26742; S00533; S05723
R:Wilson, P.G.; Culbertson, M.R.
J. Mol. Biol. 199, 559-573, 1988
A:Title: SUP12 suppressor protein of yeast. A fusion protein related to the EF-1 family
A:Reference number: S00733; MUID:88172503; PMID:3280807
A:Accession: S00733
A:Molecule type: DNA
A:Residues: 1-685 <WIL>
A:Cross-references: EMBL:P05453; EMBL:X07163; NID:94581; PIDN:CAA30155.1; PID:94582
R:Kushnirov, V.V.; Ter-Avanesyan, M.D.; Telckov, M.V.; Surguchov, A.P.; Smirnov, V.N.; I
Gene 66, 45-54, 1988
A:Title: Nucleotide sequence of the SUP2(SUP35) gene of Saccharomyces cerevisiae.
A:Reference number: JT0323; MUID:88329727; PMID:3047009
A:Accession: JT0323
A:Molecule type: DNA
A:Residues: 1-685 <KUS>
A:Cross-references: EMBL:M21129; NID:9172789; PIDN:AAA35133.1; PID:9172791
R:Murphy, L.; Harris, D.E.
submitted to the EMBL Data Library, November 1994
A:Reference number: S49768
A:Accession: S49768
A:Molecule type: DNA
A:Residues: 1-685 <MUR>
A:Cross-references: EMBL:Z46727; NID:91289283; PIDN:CAA86677.1; PID:91289287; GSPDB:GN00
R:Kikuchi, Y.; Shimatake, H.; Kikuchi, A.
EMBO J. 7, 1175-1182, 1988
A:Title: A yeast gene required for the G1-to-S transition encodes a protein containing a
A:Reference number: S00488; MUID:88296422; PMID:2841115
A:Accession: S00488
A:Molecule type: DNA
A:Residues: 1-52, 'C', 54-685 <KIK>
A:Cross-references: GB:Y00829; EMBL:Y00859; NID:93711; PIDN:CAA68760.1; PID:93712
C:Genetics:
A:Gene: SGD:SUP35; SUP12; GST1; SUP2; MIPS:YDR172w
A:Cross-references: SGD:S0002579; MIPS:YDR172w
A:Map position: 4R
C:Superfamily: suppressor 2 protein; translation elongation factor Tu homology
C:Keywords: duplication; GTP binding; nucleotide binding; P-loop; phosphoprotein; tandem
F:1-123/Domain: A <DOM1>
F:42-119/Region: 10-residue repeats
F:124-253/Domain: charged <DOM2>
F:159-222/Region: glutamic acid/lysine-rich
F:254-685/Domain: C <DOM4>
F:261-409/Domain: translation elongation factor Tu homology <ETU>
F:267-274/Region: nucleotide-binding motif A (P-loop)
F:406-409/Region: GTP-binding NKXD motif

RESULT 7
T20847
hypothetical protein F13E9.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20847
R:McMurray, A.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19332
A:Accession: T20847
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-409 <WIL>
A:Cross-references: UNIPROT:Q19414; EMBL:Z69383; PIDN:CAA93412.1; GSPDB:GN00022; CESP:F13E9.4
A:Experimental source: clone F13E9
C:Genetics:
A:Gene: CESP:F13E9.4
A:Map position: 4
A:Introns: 32/1; 275/3; 337/3
C:Superfamily: loricrin

	4	SNQGNQNTQQ--YSQNGNQOQGNNRVOGYQAYNAAQPAGGYYQN-----YOGYGSYQ	56
Qy		: :	
Db	55	TGCGLLSSYYQQGYGQNGSMQYSQ---GGYGENSQDDYGSQSQSGMVGQYGS	110
		: :	
Qy	57	QGGYQQYNPDAGYQQYNPQGGYQ-----QYNPQGGYQQQFNPQCGGRNYK	102
		: :	
Db	111	QSYGQQ-----AFACQQRPPQGGFSNFCGQASGSMNSFGCGGYGQNGFGCGSGFS	165
		: :	
Qy	103	-----NFNYNNLQGYAQGFQPSQ	123
		: :	
Db	166	GSGGWGNSLSANSNGNN-OGSSSGYO--NNOG	197
		: :	

RESULT 8
D89723
protein F39D8.1b [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D89723
P:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
A:Reference number: A75000; MURID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999;
A:Accession: D89723
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <STO>
A:Cross-references: UNIPROT.Q814J3; GB:chr_X; PIDN.CAA93661.1; PID:g3876953; GSPDB:GN000000000
C:Genetics:
A:Gene: F39D8.1b
A:Map position: X

```

Query Match      21.0%; Score 147.5; DB 2; Length 940;
Best Local Similarity 34.9%; Pred. No. 2.7e-05;
Matches 52; Conservative 7; Mismatches 63; Indels 27; Gaps 8;

Qy      1  MSDSNQGNQNTQYQYSONGQNGNRRYQGYQAYNA-QAQ---PAGYYQNY-QGYSGY 55
      :  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      222  VGDQNGNQFQDPNGQYQNGNQNTGTYGDSQNGQGPDPNGQYQNGNQNGNI 281

Qy      56  QQGGYQYQNPDAGYQQQYNPQG--GYQQYNP-----QGGYQQGFN-----PQGGRGNY 101
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      282  NAGSYVDQNGFQGGDQNGINAGGYVGNQNGNPFQFDQNGNQNGYQGYQGSLPAGSVDY 341
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```


QY		7	GNNQQ-----NYQQYSQNQNO-----QGNNRYQGY---QAYNA-----	37
Dd		573	GNNYQTSAATAGYSQFYSGNGCAGNAGGGGAGSGYSSYYQEGYNAPTPPKPFVKPPPP	632
QY		38	--QAQPAGGYQN-----YQYSGYQGGYQQYNPDAGYQQQYNPQGGYQQYNP----	85
Dd		633	QQQQQQPPPHASNPKPFSYNGYQGH-QGGQQQQPQQQQQTYN-QNQSYNYPQKQK	690
QY		86	GGYQQQFPNPGGRGNKNFNNNLQ	112
Dd		691	GGYNQGTGAASAGSY---NYSNSYT	714

Search completed: July 1, 2005, 18:21:01
Job time : 10.1876 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 18:05:25 ; Search time 19.5919 Seconds
(without alignments)
3214.888 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_123

Perfect score: 701

Sequence: 1 MSDSNQGNQNYQYQSQNG.....FNYNNNLQYQAGFPQSQG 123

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	701	100.0	123	Q8TFR6	Q8tfr6 saccharomyc
2	701	100.0	236	Q8TFA9	Q8tfa9 saccharomyc
3	701	100.0	237	Q8TFQ8	Q8tfq8 saccharomyc
4	701	100.0	240	Q8TFR0	Q8tfr0 saccharomyc
5	701	100.0	242	Q8TFR4	Q8tfr4 saccharomyc
6	701	100.0	685	1 ERP2_YEAST	P05453 saccharomyc
7	701	100.0	685	2 Q8TFB8	Q8tfb8 saccharomyc
8	701	100.0	685	2 Q8HGV1	Q8hgv1 saccharomyc
9	697	99.4	233	2 Q8TFQ9	Q8tfq9 saccharomyc
10	696	99.3	235	2 Q8TFR1	Q8tfr1 saccharomyc
11	696	99.3	243	2 Q8TFR3	Q8tfr3 saccharomyc
12	695	99.1	224	2 Q07089	Q07089 saccharomyc
13	692	98.7	251	2 Q96TL8	Q96tl8 saccharomyc
14	692	98.7	251	2 Q96TMO	Q96tmo saccharomyc
15	692	98.7	251	2 Q96UJ7	Q96uj7 saccharomyc
16	692	98.7	429	2 Q96TJ4	Q96tj4 saccharomyc
17	692	98.7	429	2 Q96UJ0	Q96uj0 saccharomyc
18	692	98.7	429	2 Q96UJ2	Q96uj2 saccharomyc
19	688	98.1	429	2 Q96UI9	Q96ui9 saccharomyc
20	687	98.0	251	2 Q96UJ4	Q96uj4 saccharomyc
21	687	98.0	251	2 Q96UJ5	Q96uj5 saccharomyc
22	687	98.0	251	2 Q96UJ6	Q96uj6 saccharomyc
23	687	98.0	251	2 Q96UJ8	Q96uj8 saccharomyc
24	687	98.0	429	2 Q96TQ9	Q96tq9 saccharomyc
25	687	98.0	429	2 Q96UJ1	Q96uj1 saccharomyc
26	686	97.9	120	2 Q8TFR9	Q8tfr9 saccharomyc
27	683	97.4	429	2 Q96UI8	Q96ui8 saccharomyc
28	665	94.9	435	2 Q6Q7I2	Q6q7i2 saccharomyc
29	665	94.9	435	2 Q6Q7I3	Q6q7i3 saccharomyc
30	665	94.9	435	2 Q6Q7I6	Q6q7i6 saccharomyc
31	660	94.2	435	2 Q6Q7I1	Q6q7i1 saccharomyc

32	660	94.2	435	2	Q6Q7I4	Q6q7i4 saccharomyc
33	625	89.2	434	2	Q96UJ3	Q96uj3 saccharomyc
34	569.5	81.2	216	2	Q8TFR2	Q8tfr2 saccharomyc
35	569.5	81.2	227	2	Q8TFR5	Q8tfr5 saccharomyc
36	569.5	81.2	666	2	Q8TFR7	Q8tfr7 saccharomyc
37	569.5	81.2	666	2	Q8TFR8	Q8tfr8 saccharomyc
38	244	34.8	688	2	Q6FVM2	Q6fvm2 candida gla
39	229.5	32.7	691	2	Q750T4	Q750t4 ashbya goss
40	229.5	32.7	712	2	Q9HG17	Q9hgi7 candida mal
41	228	32.5	157	2	Q8TFQ1	Q8tfq1 candida alb
42	228	32.5	192	2	Q8TFP9	Q8tfp9 candida alb
43	228	32.5	299	2	Q9UVK3	Q9uvk3 candida alb
44	228	32.5	715	1	ERP2_CANAL	O13354 candida alb
45	223	31.8	174	2	Q8TFQ0	Q8tfq0 candida alb

ALIGNMENTS

RESULT 1

Q8TFR6 PRELIMINARY; PRT; 123 AA.
AC Q8TFR6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Prion protein (Fragment).
GN Name=SUP35;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC15;
RX MEDLINE=22773310; PubMed=12890024;
RA Resende C.G., Outeiro T.F., Sands L., Lindquist S., Tuite M.F.;
RT "Prion protein gene polymorphisms in Saccharomyces cerevisiae.";
RL Mol. Microbiol. 49:1005-1017(2003).
DR EMBL; AY028648; AAK26179.1; -.
KW Prion.
FT NON TER 123 123
SQ SEQUENCE 123 AA; 14048 MW; C1BA73BCFC4997A7 CRC64;

Query Match 100.0%; Score 701; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.5e-50;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDSNQGNQNYQYQSQNGNQGNQNNRYQGYQAYNAQAQAGGYQNYQYSGYQSGY 60
DB 1 MSDSNQGNQNYQYQSQNGNQGNQNNRYQGYQAYNAQAQAGGYQNYQYSGYQSGY 60
QY 61 QQYNPDAGYQQYNPQGGYQQYNPQGGYQQQFNQGGYKFNFNNNLQYQAGFPQ 120
DB 61 QQYNPDAGYQQYNPQGGYQQYNPQGGYQQQFNQGGYKFNFNNNLQYQAGFPQ 120
QY 121 SQG 123
DB 121 SQG 123

RESULT 2

Q8TFA9 PRELIMINARY; PRT; 236 AA.
AC Q8TFA9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Prion protein (Fragment).
GN Name=SUP35;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

```
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC110, and SC113;
RA MEDLINE=22773310; PubMed=12890024;
RX Resende C.G., Outeiro T.F., Sands L., Lindquist S., Tuite M.F.;
RT "Prion protein gene polymorphisms in Saccharomyces cerevisiae.";
RL Mol. Microbiol. 49:1005-1017(2003).
DR EMBL; AY028653; AAK26184.1; -.
DR EMBL; AY028656; AAK26187.1; -.
KW Prion.
FT NON_TER 236 236
SQ SEQUENCE 236 AA; 26621 MW; 698FD2AA89CBB471 CRC64;

Query Match 100.0%; Score 701; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 2.9e-50;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSDSNQGNQNYQYYSQNGNQGNRYQGYQAYNAQAQAGGYQNYQYSGYQGGY 60
Db 1 MSDSNQGNQNYQYYSQNGNQGNRYQGYQAYNAQAQAGGYQNYQYSGYQGGY 60

Qy 61 QQYNPDAGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGY 120
Db 61 QQYNPDAGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGY 120

Qy 121 SQG 123
Db 121 SQG 123

RESULT 3
Q8TFQ8 PRELIMINARY; PRT; 237 AA.
AC Q8TFQ8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Prion protein (Fragment).
GN Name=SUP35;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC115;
RA MEDLINE=22773310; PubMed=12890024;
RX Resende C.G., Outeiro T.F., Sands L., Lindquist S., Tuite M.F.;
RT "Prion protein gene polymorphisms in Saccharomyces cerevisiae.";
RL Mol. Microbiol. 49:1005-1017(2003).
DR EMBL; AY028659; AAK26190.1; -.
KW Prion.
FT NON_TER 237 237
SQ SEQUENCE 237 AA; 26692 MW; 2D698FD2AA89CBB4 CRC64;

Query Match 100.0%; Score 701; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 2.9e-50;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSDSNQGNQNYQYYSQNGNQGNRYQGYQAYNAQAQAGGYQNYQYSGYQGGY 60
Db 1 MSDSNQGNQNYQYYSQNGNQGNRYQGYQAYNAQAQAGGYQNYQYSGYQGGY 60

Qy 61 QQYNPDAGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGY 120
Db 61 QQYNPDAGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGY 120

Qy 121 SQG 123
Db 121 SQG 123

RESULT 4
Q8TFQ8 PRELIMINARY; PRT; 240 AA.
AC Q8TFQ8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Prion protein (Fragment).
GN Name=SUP35;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC114;
RA MEDLINE=22773310; PubMed=12890024;
RX Resende C.G., Outeiro T.F., Sands L., Lindquist S., Tuite M.F.;
RT "Prion protein gene polymorphisms in Saccharomyces cerevisiae.";
RL Mol. Microbiol. 49:1005-1017(2003).
DR EMBL; AY028657; AAK26188.1; -.
KW Prion.
FT NON_TER 240 240
SQ SEQUENCE 240 AA; 27046 MW; B7BFB4AD698FD2AA CRC64;

Query Match 100.0%; Score 701; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSDSNQGNQNYQYYSQNGNQGNRYQGYQAYNAQAQAGGYQNYQYSGYQGGY 60
Db 1 MSDSNQGNQNYQYYSQNGNQGNRYQGYQAYNAQAQAGGYQNYQYSGYQGGY 60

Qy 61 QQYNPDAGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGY 120
Db 61 QQYNPDAGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGYQQYYPGGY 120

Qy 121 SQG 123
Db 121 SQG 123

RESULT 5
Q8TFQ8 PRELIMINARY; PRT; 242 AA.
AC Q8TFQ8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Prion protein (Fragment).
GN Name=SUP35;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC18;
RA MEDLINE=22773310; PubMed=12890024;
RX Resende C.G., Outeiro T.F., Sands L., Lindquist S., Tuite M.F.;
RT "Prion protein gene polymorphisms in Saccharomyces cerevisiae.";
RL Mol. Microbiol. 49:1005-1017(2003).
DR EMBL; AY028651; AAK26182.1; -.
KW Prion.
FT NON_TER 242 242
SQ SEQUENCE 242 AA; 27303 MW; DA5B27BBF4AD698F CRC64;

Query Match 100.0%; Score 701; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSDSNQGNQNYQYYSQNGNQGNRYQGYQAYNAQAQAGGYQNYQYSGYQGGY 60
```



```

SQ  SEQUENCE      685 AA;  76551 MW;  43912A6D77DFA153  CRC64;

Query Match      100.0%;  Score 701;  DB 1;  Length 685;
Best Local Similarity 100.0%;  Pred. No. 8.8e-50;
Matches 123;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      1  MSDSNQGNQNYQYQYSGNGQQGNRRYQGYQAYNAQAQAGGYQNYQYSGYQGGY  60
      |||||
Db      1  MSDSNQGNQNYQYQYSGNGQQGNRRYQGYQAYNAQAQAGGYQNYQYSGYQGGY  60
      |||||

QY      61  QQYNPDAGYQQQYNPQGGYQOQYNPQGGYQQQFPNPGGGRNFKNFNNNLLQGYQAGFPQ  120
      |||||
Db      61  QQYNPDAGYQQQYNPQGGYQOQYNPQGGYQQQFPNPGGGRNFKNFNNNLLQGYQAGFPQ  120
      |||||

QY      121  SQG 123
      |||
Db      121  SQG 123
      |||

RESULT 7
ID      Q8TFB8      PRELIMINARY;      PRT;      685 AA.
AC      Q8TFB8;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Prion protein.
GN      Name=SUP35;
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX      NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SCI2, and SCI6;
RX      MEDLINE=22773310; PubMed=12890024;
RT      Resende C.G., Outeiro T.F., Sands L., Lindquist S., Tuite M.F.;
RT      "Prion protein gene polymorphisms in Saccharomyces cerevisiae.";
RL      Mol. Microbiol. 49:1005-1017(2003).
DR      EMBL; AY028645; AAK26176.1; -.
DR      EMBL; AY028649; AAK26180.1; -.
DR      SGD; S000002579; SUP35.
DR      GO; GO:0005525; F:GTP binding; IEA.
DR      GO; GO:0003747; F:translation release factor activity; IEA.
DR      GO; GO:0006412; P:protein biosynthesis; IEA.
DR      GO; GO:0006415; P:translational termination; IEA.
DR      InterPro; IPRO04160; EFTU Cterm.
DR      InterPro; IPRO04161; EFTU D2.
DR      InterPro; IPRO09001; Elong init C.
DR      InterPro; IPRO00795; ProtSyn GTPbind.
DR      InterPro; IPRO09000; Translat_factor.
DR      InterPro; IPRO03285; Yeast_ERF.
DR      Pfam; PF00009; GTP EFTU; 1.
DR      Pfam; PF03144; GTP EFTU D2; 1.
DR      Pfam; PF03143; GTP EFTU D3; 1.
DR      PRINTS; PR00315; ELONGATNFCT.
DR      PRINTS; PR01343; YEASTERF.
DR      PROSITE; PS00301; EFATOR GTP; UNKNOWN 1.
DR      GTP-binding; Prion; Protein biosynthesis.
SQ  SEQUENCE      685 AA;  76609 MW;  43912A6DDBAF4B53  CRC64;

```

	Query Match	100.0%;	Score 701;	DB 2;	Length 685;
	Best Local Similarity	100.0%;	Pred. NO. 8.e-50;		
	Matches 123;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MSDSNGNNQNNYYQYSSQNGNQGGNNRYYGYQAYNAQAQAGYGYQNYQYSGYQGGY	60		
Db	1	MSDSNGNNQNNYYQYSSQNGNQGGNNRYYGYQAYNAQAQAGYGYQNYQYSGYQGGY	60		
QY	61	QQYNPDAGYQQQYNPDGGYQQYNPDGGYQQQFNPDGGRGNYKNFNYYNNNLOGYQAGFPQ	120		
Db	61	QQYNPDAGYQQQYNPDGGYQQYNPDGGYQQQFNPDGGRGNYKNFNYYNNNLOGYQAGFPQ	120		

Qy	121	SQG	123
Db	121	SQG	123
RESULT 8			
Q9HGV1		PRELIMINARY;	PRT; 685 AA.
ID	Q9HGV1		
AC	Q9HGV1;		
DT	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	SUP35 allouppressor mutant sal3-4.		
DE	Name=SUP35;		
GN	Saccharomyces cerevisiae (Baker's yeast).		
OS	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
NCBI_TaxID=4932;			
CX	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=BSC783/4a sal3-4;		
KC	Resende C.G., Duarte J.B., Tuite M.F.;		
RA	Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.		
RL	EMBL; AF263099; AAF9684.1; -		
DR	GO; GO:0005525; F:GTP binding; IEA.		
DR	GO; GO:0003747; F:translation release factor activity; IEA.		
DR	GO; GO:0006412; P:protein biosynthesis; IEA.		
DR	GO; GO:0006415; P:translational termination; IEA.		
DR	InterPro; IPR004160; EFTU Cterm.		
DR	InterPro; IPR004161; EFTU D2.		
DR	InterPro; IPR009001; Elong init C.		
DR	InterPro; IPR000795; ProtSyn.GTPbind.		
DR	InterPro; IPR009000; Translat factor.		
DR	InterPro; IPR003285; Yeast_ERF.		
DR	Pfam; PF00009; GTP_EFTU; 1.		
DR	Pfam; PF03144; GTP_EFTU_D2; 1.		
DR	Pfam; PF03143; GTP_EFTU_D3; 1.		
DR	PRINTS; PR00315; ELONGATNCT.		
DR	PRINTS; PR01343; YEASTERF.		
DR	PROSITE; PS00301; EFACITOR_GTP; UNKNOWN_1.		
KW	GTP-binding; Protein biosynthesis.		
SW	SEQUENCE 685 AA; 76550 MW; F39S8EC37D750154 CRC64;		
SEQ			
Query Match 100.0%; Score 701; DB 2; Length 685;			
Best Local Similarity 100.0%; Pred.No.8.e-50;			
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MDSNGNNQOYQOYSQNGNQGGNRYYGYQAYNAQAQAGGYQNYQGYSGYQGGY	60
Db	1	MDSNGNNQOYQOYSQNGNQGGNRYYGYQAYNAQAQAGGYQNYQGYSGYQGGY	60
Qy	61	QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQQFPQGGRGNYKNFNANNLNQGYAQGFQPQ	120
Db	61	QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQQFPQGGRGNYKNFNANNLNQGYAQGFQPQ	120
Qy	121	SQG	123
Db	121	SQG	123
RESULT 9			
Q8TFQ9		PRELIMINARY;	PRT; 233 AA.
ID	Q8TFQ9		
AC	Q8TFQ9;		
DT	01-JUN-2002 (TrEMBLrel. 21, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)		
DE	Prión protein (Fragment).		
DE	Name=SUP35;		
GN	Saccharomyces cerevisiae (Baker's yeast).		
OS	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
NCBI_TaxID=4932;			

Search completed: July 1, 2005, 18:19:52
Job time : 20.5919 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 18:17:24 ; Search time 9.04241 Seconds
(without alignments)
1015.419 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_123

Perfect score: 701

Sequence: 1 MSDSNQGNQYQYQSQNG.....FNYNNLQGYQAGFPQSQG 123

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pap:*
 - 2: /cgn2_6/ptodata/1/iaa/5B COMB.pap:*
 - 3: /cgn2_6/ptodata/1/iaa/6A COMB.pap:*
 - 4: /cgn2_6/ptodata/1/iaa/6B COMB.pap:*
 - 5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pap:*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	148.5	21.2	207	4	US-10-029-180-70 Sequence 70, Appl
2	141	20.1	391	4	US-09-949-016-6646 Sequence 6646, Ap
3	141	20.1	392	4	US-09-949-016-10150 Sequence 10150, A
4	136	19.4	462	4	US-09-919-039-324 Sequence 324, App
5	136	19.4	526	4	US-09-538-092-1080 Sequence 1080, Ap
6	129	18.4	253	4	US-09-248-796A-19227 Sequence 19227, A
7	128	18.3	907	4	US-09-949-016-9750 Sequence 9750, Ap
8	128	18.3	907	4	US-09-949-016-9751 Sequence 9751, Ap
9	127	18.1	494	4	US-09-248-796A-14156 Sequence 14156, A
10	127	18.1	1196	4	US-09-248-796A-19876 Sequence 19876, A
11	124	17.7	159	4	US-09-248-796A-19824 Sequence 19824, A
12	122.5	17.5	284	4	US-09-902-540-12642 Sequence 12642, A
13	122	17.4	210	4	US-09-248-796A-25835 Sequence 25835, A
14	121	17.3	485	2	US-08-749-391-2 Sequence 2, Appli
15	121	17.3	485	3	US-09-390-200-2 Sequence 2, Appli
16	121	17.3	954	4	US-09-570-856B-31 Sequence 31, Appl
17	120	17.1	180	4	US-10-029-180-38 Sequence 38, Appl
18	120	17.1	286	4	US-09-248-796A-16992 Sequence 16992, A
19	118.5	16.9	456	4	US-09-248-796A-15862 Sequence 15862, A
20	117	16.7	579	3	US-09-521-780-5 Sequence 5, Appli
21	117	16.7	579	3	US-09-521-780-6 Sequence 6, Appli
22	116	16.5	298	3	US-08-122-458D-11 Sequence 11, Appl
23	116	16.5	588	4	US-09-248-796A-20839 Sequence 20839, A
24	112.5	16.0	188	4	US-09-248-796A-21795 Sequence 21795, A
25	112.5	16.0	1897	4	US-09-792-024-98 Sequence 98, Appl
26	112	16.0	302	4	US-10-029-180-18 Sequence 18, Appl
27	111.5	15.9	296	2	US-08-700-637-4 Sequence 4, Appli

28	110.5	15.8	161	4	US-09-270-767-42771 Sequence 42771, A
29	110.5	15.8	1179	4	US-09-949-016-7088 Sequence 7088, Ap
30	110	15.7	909	4	US-09-248-796A-16165 Sequence 16165, A
31	109	15.5	521	4	US-09-270-767-44481 Sequence 44481, A
32	108	15.4	229	4	US-09-248-796A-24831 Sequence 24831, A
33	107	15.3	378	4	US-10-164-595-2 Sequence 2, Appli
34	105.5	15.0	228	4	US-10-029-180-16 Sequence 16, Appl
35	105.5	15.0	712	4	US-09-248-796A-19645 Sequence 19645, A
36	104.5	14.9	170	4	US-09-270-767-41268 Sequence 41268, A
37	104.5	14.9	170	4	US-09-270-767-56484 Sequence 56484, A
38	104.5	14.9	657	4	US-09-248-796A-19232 Sequence 19232, A
39	104	14.8	307	1	US-07-982-112-2 Sequence 2, Appli
40	104	14.8	536	4	US-09-270-767-43766 Sequence 43766, A
41	103.5	14.8	78	4	US-09-328-352-6915 Sequence 6915, Ap
42	103.5	14.8	362	1	US-08-437-027-21 Sequence 21, Appl
43	103.5	14.8	365	1	US-08-437-027-20 Sequence 20, Appl
44	103.5	14.8	591	4	US-09-949-016-10914 Sequence 10914, A
45	103.5	14.8	591	4	US-09-949-016-10915 Sequence 10915, A

ALIGNMENTS

RESULT 1
US-10-029-180-70
; Sequence 70, Application US/10029180
; Patent No. 6806082
; GENERAL INFORMATION:
; APPLICANT: Cali, Brian M.
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin T.
; APPLICANT: Milna, G. Todd
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeffry C.
; APPLICANT: Trueheart, Josh
; APPLICANT: Zhang, Lixin
; TITLE OF INVENTION: No. 6806082el Regulators of Fungal Gene Expression
; FILE REFERENCE: MIC-004
; CURRENT APPLICATION NUMBER: US/10/029,180
; CURRENT FILING DATE: 2001-12-22
; PRIOR APPLICATION NUMBER: US 60/257,431
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fungal gene
US-10-029-180-70

Query Match	21.2%	Score 148.5;	DB 4;	Length 207;
Best Local Similarity	36.5%	Pred No. 2.5e-07;		
Matches	46;	Conservative	8;	Mismatches 47;
			Indels	25;
			Gaps	8;
QY	13	YQYYSQNGNQGNRYQGYQAYNAQAQAPAGY--YQNYQGYSGYQGGYQGY-NPDAGY	69	
Db	4	YDQYNGGYGQGGYQGY-QGY-----GQPGYQPGYGGQGGHQQQYQGPQHG	56	
QY	70	-QQYNPQGGYQGY-----NPGGYQQQFNPGGGRNY-----KVFYNNNLQGYQAGF	117	
Db	57	GQGGYQGGSGSDYYAGQQHQQQGYGQ---QGSSDYAGQQHQHQHGGHQQDNRQGGY	113	
QY	118	QPOSQG 123		
Db	114	EQQHG 119		

RESULT 2
US-09-949-016-6646
; Sequence 6646, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6646
 ; LENGTH: 391
 ; TYPE: PR1
 ; ORGANISM: Human
 ; US-09-949-016-6646

Query Match 20.1%; Score 141; DB 4; Length 391;
 Best Local Similarity 35.7%; Pred. No. 2.9e-06;
 Matches 50; Conservative 9; Mismatches 55; Indels 26; Gaps 10;
 QY 1 MSDSNQNNQNYQYQY-----SONGNQQ--GNNRYQGYQ-AYNAQAQAPAG---GYTQNY 49
 DB 239 MGQVQNQHMMGQRQ1PPYRPPQGPQQYSGQEDYGDYSHGGQGPPEGMMNQYYPDG 298
 QY 50 QGYSGYQGGYQYQYNDAGY---QQYNPQGGYQYQYNDAGYQYQYQYQYQYQYQYQYQY 104
 DB 299 NSQYGGQQDDAYGGPPPGQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 352
 QY 105 NYNNLQGYQ-AGFQPOSQ 123
 DB 353 PQG---GQYQYGYRPTQPG 369

RESULT 3
 US-09-949-016-10150
 ; Sequence 10150, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10150
 ; LENGTH: 392
 ; TYPE: PR1
 ; ORGANISM: Human
 ; US-09-949-016-10150

Query Match 20.1%; Score 141; DB 4; Length 392;
 Best Local Similarity 35.7%; Pred. No. 2.9e-06;
 Matches 50; Conservative 9; Mismatches 55; Indels 26; Gaps 10;
 QY 1 MSDSNQNNQNYQYQY-----SONGNQQ--GNNRYQGYQ-AYNAQAQAPAG---GYTQNY 49
 DB 240 MGQVQNQHMMGQRQ1PPYRPPQGPQQYSGQEDYGDYSHGGQGPPEGMMNQYYPDG 299
 QY 50 QGYSGYQGGYQYQYNDAGY---QQYNPQGGYQYQYNDAGYQYQYQYQYQYQYQYQYQY 104

Db 300 NSQYGGQQDDAYGGPPPGQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 353
 QY 105 NYNNLQGYQ-AGFQPOSQ 123
 DB 354 PQG---GQYQYGYRPTQPG 370
 RESULT 4
 US-09-919-039-324
 ; Sequence 324, Application US/09919039
 ; Patent No. 6727066
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaser, Matthew R.
 ; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
 ; FILE REFERENCE: PA-0035 US
 ; CURRENT APPLICATION NUMBER: US/09/919,039
 ; PRIOR FILING DATE: 2002-09-09
 ; PRIOR APPLICATION NUMBER: 60/222,113
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 401
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 324
 ; LENGTH: 462
 ; TYPE: PR1
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. 6727066 1813444CD1
 ; US-09-919-039-324

Query Match 19.4%; Score 136; DB 4; Length 462;
 Best Local Similarity 27.0%; Pred. No. 1.1e-05;
 Matches 51; Conservative 9; Mismatches 57; Indels 72; Gaps 10;
 QY 4 SNOGNQNNQY 45
 DB 29 SSQY 88
 QY 46 --YQNYQY 146
 DB 89 SSYQY 78
 QY 79 YQY 114
 DB 147 -QSYNPPQY 205
 QY 115 AGFQPOSQ 123
 DB 206 GGYQY 214

RESULT 5
 US-09-538-092-1080
 ; Sequence 1080, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Giot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CurPatSeqformatter Version 0.9
 ; SEQ ID NO 1080
 ; LENGTH: 526
 ; TYPE: PR1
 ; ORGANISM: Homo sapiens
 ; FEATURE:

```
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P35637
US-09-538-092-1080

Query Match      19.4%; Score 136; DB 4; Length 526;
Best Local Similarity 27.0%; Pred. No. 1.3e-05;
Matches 51; Conservative 9; Mismatches 57; Indels 72; Gaps 10;

QY   4  SNQGNNOQTQQYSONGNQ-QGNNRYQGY-----QAYNAQAQP-----AGGY----- 45
Db   29  SSQPYGGQSYGYSGSDTSTGYSQSSYSYCQSQNTGTGTPTQGVTGGYGSSOSSQ 88
QY   46  --YQNYQGYGY-----QCGGYQOYNPDAGYQOOQYNPQGG 78
Db   89  SSYGQQQSSYPGYGQAPSPSTGSTGYGSSSSSSYSYGPQPSGSYSQQPSYGGQQSYGQQ-- 146
QY   79  YQOYNPOGGY--QQQFNPQ-----GGRGNY-----KNFNYYNNNLQ---GYQ 114
Db   147 -QSTNPPQGYGQQNQYNSSSGGGGGGGGGNYGDQDSSMSGGGGGYGNQDQSGGGGS 205
QY   115 AGFOPOQSQG 123
Db   206 GGYGQQDRG 214

RESULT 6
US-09-248-796A-19227
; Sequence 19227, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19227
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (234), (249)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-19227

Query Match      18.4%; Score 129; DB 4; Length 253;
Best Local Similarity 34.2%; Pred. No. 2.7e-05;
Matches 39; Conservative 10; Mismatches 35; Indels 30; Gaps 7;

QY   11  QNYQOVSQNG--NQCGNNRYQGYQAYNAQAQPAQYQNYQGYSGYQCGGYQOYNPDA 67
Db   152 RSVLYQLTNHHPDHQLQNQSYLSY-----CQRPYGGGSGYNOQGY--YNOQG 198
QY   68  GYOQQQ----YNPQGGY----QQYNPOGGYQQQFPNPGQGRGNKFNFY--NNNLOGY 113
Db   199 RYNOQGNRYNOQGRYSQQSYQVQYGSNTQRFN-----NNQXYNQSSNNRSRGY 246

RESULT 7
US-09-949-016-9750
; Sequence 9750, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```



```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,391
FILING DATE: 13-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Donna M. Ferber
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 93-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-749-391-2

Query Match 17.3%; Score 121; DB 2; Length 485;
Best Local Similarity 35.5%; Pred. No. 0.00036;
Matches 49; Conservative 8; Mismatches 55; Indels 26; Gaps 8;

QY 6 QGNNQNYQYSSNGNQCCNNRYQGYQAYNAQ-----AQAGGYYQNY-QGYSGYQQGGY 60
Db 304 QGNDQGG-QQPPQGGQPPQGNDDQQGQQPPQPGQPGCGNPGGSDFNWNG--GSPWGGN 360
QY 61 QQYNPDAGYQQQYNPQGGYQQYNPQGGYQ-----QQFNPOGGR-----GNYKNF 104
Db 361 QGGSPPWG-GNOGPNPWSGNGSGSPWGGNGQGSPPWGGNQGNPWSGNGQGSPPWGGNQGN 419
QY 105 NYNNNLQGY-QAGFPQOS 121
Db 420 PWGGNQWGAQNAAPQS 437

RESULT 15
US-09-390-200-2
Sequence 2, Application US/09390200
Patent No. 6137032
GENERAL INFORMATION:
APPLICANT: Cheng, Kuo-Joan
APPLICANT: Selinger, Leonard B.
APPLICANT: Liu, Jin-Hao
APPLICANT: Hu, Youji
APPLICANT: Forsberg, Cecil W.
APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: A Xylanase Obtained From an
NUMBER OF INVENTION: Anaerobic Fungus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/390,200
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US/08/749,391

```

```
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna M. Feiber
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 93-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-390-200-2

Query Match      17.3%; Score 121; DB 3; Length 485;
Best Local Similarity 35.5%; Pred.No. 0.00036;
Matches 49; Conservative 8; Mismatches 55; Indels 26; Gaps 8;

Qy      6 QGNNQQNYQQYSQNGNQOQGNRRYQGYQAYNAQ-----AQPAGGYQNY-QGYSGYQGGY 60
Db     304 QGNPQGG-QQPPQGGQPPQGNDDQGGQPPQPPQGGNPGGSDFNWNNQG--GSPWGEN 360

Qy     61 QQYNPDAGYQQQYNPQGGYQQYNPQGGYQ-----QQFNPQGGR-----GNYKNF 104
Db    361 QGGSFWGG-NQGGNPNWGGNQGGSPWGGNQGGSPWGGNQGGNPGWGGNQGGSPWGGNQGGN 419

Qy    105 NYNNLQGY-QAGFPQPS 121
Db    420 PWGNGWGAQNAAAPQS 437
```

Search completed: July 1, 2005, 18:39:40
Job time : 10.0424 secs


```
RESULT 2
US-10-032-585-7274
; Sequence 7274, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7274
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7274

Query Match      31.8%; Score 223; DB 14; Length 721;
Best Local Similarity 43.4%; Pred. No. 7.2e-14;
Matches 59; Conservative 11; Mismatches 42; Indels 24; Gaps 9;

QY      8 NNQNNYQQY-----SONGN-----QQGNNRYQGYQAYNAQAQ-----PAGGY--YQNYQ 50
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2 SDQNTQDQLSCAMANASLNGDSKQQQQQQQQQQQQQQNNYNNPNAQSFVPQGGYQQFQQF 61

QY      51 GYSGYQQ--GGYQQNP--DAGYQQNP--DAGYQQNP--DAGYQQNP--DAGYQQNP--DAGYQQNP 107
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      62 PQQQQQQQGGYQNYQYQGGYQQYQNNRGGYQQGNNRGGYQQYNN--NRGGYQYNNQ 118

QY      108 NNLCYQA--GFQPSQ 122
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      119 QQYGGYQQYNSQPQQ 134

RESULT 3
US-10-732-923-17479
; Sequence 17479, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 17479
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Debaryomyces hansenii
US-10-732-923-17479

Query Match      27.7%; Score 194.5; DB 17; Length 701;
Best Local Similarity 38.5%; Pred. No. 5e-11;
Matches 60; Conservative 8; Mismatches 31; Indels 57; Gaps 12;

QY      1 MSDSNQGN-----NQNNYQQYSONGNGQQGNNRYQGYQAYNAQAQPA 42
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MSDDQQYNNQDKLQDFQNTSIGSGEQQQQYQYQQ-----QPQNN-----FNANSAPT 50

QY      43 -----GGYQNYQ--GYSGYQQGYQQYNNPDAGYQQYNNPDAGYQQY--NPQGGY 88
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      51 FTSGQQGGYQQGYQGGYQNYSGGYQYNN--QGY-QYNNQ--GYQYQNNRGGY 105

QY      89 QQQFNPQGGRGNNYNNNNLNQYQAGFQP--QSQG 123
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db      106 -NSYNNRGYNNYNNYNN-----QQDQQPQVQNG 132

RESULT 4
US-10-732-923-17478
; Sequence 17478, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 17478
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Zygosaccharomyces rouxii
US-10-732-923-17478

Query Match      27.1%; Score 190; DB 17; Length 662;
Best Local Similarity 46.8%; Pred. No. 1.3e-10;
Matches 58; Conservative 1; Mismatches 39; Indels 26; Gaps 10;

QY      1 MSDSNQGNQQYQQYSONGNGQQGNNRYQGY-QAYNAQAQAGGYQYNNYQGYSGYQQGG 59
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MSDPNQ--NGQ-----QGGQQNAGNNYQQYFQKLTQQAQ--AGGYQPYGGYGGY--GG 49

QY      60 YQYNNPDAGYQQYNNPQGGYQ--QYNNPQGGYQQQFNPQGGRGNNYNNNNLNQYQAGFQ 118
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      50 YGYQYQYGGYQQFY--QDQQQAQQGAYNGYPIQ--AQQAPGGFN--NYNNQ-----FQ 96

QY      119 PQSQ 122
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      97 PQQQ 100

RESULT 5
US-10-732-923-17477
; Sequence 17477, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 17477
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Yarrowia lipolytica
US-10-732-923-17477

Query Match      25.0%; Score 175; DB 17; Length 742;
Best Local Similarity 35.1%; Pred. No. 4.9e-09;
Matches 59; Conservative 7; Mismatches 44; Indels 58; Gaps 11;

QY      1 MSDS--NQ--NNQNNYQQYVQ--NGNQQQGNNRYQGYQAYNAQAQ----- 40
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MSDQFNQDFANKANINEGGQGGQGGQGG-----QYGGYQAPQFVPQGSFVPQGSFVP 56

QY      41 ---PAGGYQYNNYQYSGYQ-----QGGYQQYNNP-----DAGYQQYNNP 75
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      57 SFAQGGYDQ--YQGGYQGGYQGGYQGGYQGGYQGGYQGGYQGGYQGGYQGGYQGGYQGGY 115

QY      76 QGGYQQYNNPQGGYQQQFNPQGG--RGNNYNNNNNNLNQYQAGFQPSQ 122
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      116 QGGYQ--GYQGGYDQYNGNQGGYQGGY-----DNQYGYDQPEPQEK 156
```

```

RESULT 8
US-09-866-557A-5
; Sequence 5, Application US/09866557A
; Patent No. US20020162126A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, S.
; APPLICANT: Hannon, G.
; APPLICANT: Beach, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RNA INTERFERENCE
; FILE REFERENCE: GNCA-P02-007
; CURRENT APPLICATION NUMBER: US/09/866,557A
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/189,739
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/243,097
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-866-557A-5

Query Match      23.3%   Score 163;   DB 9;   Length 1145;
Best Local Similarity 34.2%;   Pred.No. 1.3e-07;
Matches 52;   Conservative 10;   Mismatches 56;   Indels 34;   Gaps 7;

Qy      6      QGNQQ-----NYQY---SSNGNQOQGNRRYQGYQAYNAQAQAPAGGYQNYQY--- 52
Db      119    QGHHQGRGQDGGYQRPFGQGGHQQGRQGGYQRPFGQGGHQQGRQGG 178
Qy      53      -----SGYQGGYQO---YNPDAGYQO---YNPQGGYQO---YNPQGGYQO-----Q 91
Db      179    YQRPSPGQQQGGHQQGRQGGYQRPFGQGGHQQGRQGGYQRPSPGQQQGGHQQ 238
Qy      92      FNPQGGRGKFNFNNNLNQYQAGFQPSQG 123
Db      239    QGRQGGEGGYQRPSPGQQQGGHQQGRQGG 270

RESULT 9
US-09-858-862-5
; Sequence 5, Application US/09858862
; Publication No. US20040018999A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, S.
; APPLICANT: Hannon, G.
; APPLICANT: Beach, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RNA INTERFERENCE
; FILE REFERENCE: GNCA-P01-007
; CURRENT APPLICATION NUMBER: US/09/858,862
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: PCT/US01/08435
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/243,097
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/189,739
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-858-862-5

Query Match      23.3%   Score 163;   DB 11;   Length 1145;
Best Local Similarity 34.2%;   Pred.No. 1.3e-07;
Matches 52;   Conservative 10;   Mismatches 56;   Indels 34;   Gaps 7;

Qy      6      QGNQQ-----NYQY---SSNGNQOQGNRRYQGYQAYNAQAQAPAGGYQNYQY--- 52

```

```
Db 119 QGHHQGRQGGYQQRPPGQGGHQQGRQGGYQQRPPGQGGHQQGRQGG 178
QY 53 -----SGYQGGYQ--YNPDAGYQ--YNPDAGYQ--YNPDAGYQ--YNPDAGYQ--Q 91
Db 179 YQRRPSSGQGGHQQGRQGGYQQRPPGQGGHQQGRQGGYQQRPPGQGGHQQGRQGG 238
QY 92 FNPQGGGNYKFNFNNNLNQYQAGFPQSQ 123
Db 239 QGRQGGYQQRPPGQGGHQQGRQGG 270
```

RESULT 10

```
US-10-055-797-5
; Sequence 5, Application US/10055797
; Publication No. US20030084471A1
; GENERAL INFORMATION:
; APPLICANT: Hammond et al.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RNA INTERFERENCE
; FILE REFERENCE: GNCA-P03-007
; CURRENT APPLICATION NUMBER: US/10/055,797
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/189,739
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-055-797-5
```

```
Query Match 23.3%; Score 163; DB 14; Length 1145;
Best Local Similarity 34.2%; Pred. No. 1.3e-07;
Matches 52; Conservative 10; Mismatches 56; Indels 34; Gaps 7;

QY 6 QGNNQ-----NYQY-----SONGNOQGNRRYQGYQAYNAQAQAPAGGYQNYQY--- 52
Db 119 QGHHQGRQGGYQQRPPGQGGHQQGRQGGYQQRPPGQGGHQQGRQGG 178
QY 53 -----SGYQGGYQ--YNPDAGYQ--YNPDAGYQ--YNPDAGYQ--YNPDAGYQ--Q 91
Db 179 YQRRPSSGQGGHQQGRQGGYQQRPPGQGGHQQGRQGGYQQRPPGQGGHQQGRQGG 238
QY 92 FNPQGGGNYKFNFNNNLNQYQAGFPQSQ 123
Db 239 QGRQGGYQQRPPGQGGHQQGRQGG 270
```

RESULT 11

```
US-10-350-798-5
; Sequence 5, Application US/10350798
; Publication No. US2004008684A1
; GENERAL INFORMATION:
; APPLICANT: Hammond et al.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RNA INTERFERENCE
; FILE REFERENCE: CSHL-P04-010
; CURRENT APPLICATION NUMBER: US/10/350,798
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: 60/243,097
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 09/866,557
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/858,862
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: PCT/US01/08435
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/243,097
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/189,739
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 5
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-350-798-5
```

```
Query Match 23.3%; Score 163; DB 15; Length 1145;
Best Local Similarity 34.2%; Pred. No. 1.3e-07;
Matches 52; Conservative 10; Mismatches 56; Indels 34; Gaps 7;

QY 6 QGNNQ-----NYQY-----SONGNOQGNRRYQGYQAYNAQAQAPAGGYQNYQY--- 52
Db 119 QGHHQGRQGGYQQRPPGQGGHQQGRQGGYQQRPPGQGGHQQGRQGG 178
QY 53 -----SGYQGGYQ--YNPDAGYQ--YNPDAGYQ--YNPDAGYQ--YNPDAGYQ--Q 91
Db 179 YQRRPSSGQGGHQQGRQGGYQQRPPGQGGHQQGRQGGYQQRPPGQGGHQQGRQGG 238
QY 92 FNPQGGGNYKFNFNNNLNQYQAGFPQSQ 123
Db 239 QGRQGGYQQRPPGQGGHQQGRQGG 270
```

RESULT 12

```
US-10-437-963-109378
; Sequence 109378, Application US/10437963
; Publication No. US2004012334A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109378
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13541C.1.pap
US-10-437-963-109378
```

```
Query Match 21.9%; Score 153.5; DB 16; Length 671;
Best Local Similarity 34.4%; Pred. No. 6.3e-07;
Matches 42; Conservative 9; Mismatches 50; Indels 21; Gaps 5;

QY 6 QGNNQNTQYYSQNGNQGNRRYQGYQAYNAQAQAPAGGYQNYQYSGY-QGGYQYV 64
Db 550 RGMGGYGGYGGYGNNGYQ-----GGYDNGGYGGYDNGGYGGYD 596
QY 65 PDAGY---QQYNPDAGYQ--YNPDAGYQ--YNPDAGYQ--YNPDAGYQ--Q 121
Db 597 NQGGYGGYGGYGN-QGRYGNQNGYGNRGRGRGRGNW---NYRGGYRGRGGYGG 652
QY 122 QG 123
Db 653 RG 654
```

RESULT 13

```
US-10-156-761-14380
; Sequence 14380, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 18:05:05 ; Search time 70.5825 Seconds
(without alignments)
1386.327 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_253

Perfect score: 1348

Sequence: 1 MSDSNQGNQYQYQSQNG.....SADALIKQBEVDDEVND 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002a.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1348	100.0	685	AAB30792	Amino aci
2	1348	100.0	685	ABR53107	Protein s
3	1348	100.0	685	ADK63022	Disease t
4	1329	98.6	271	AAB30800	A modifie
5	1097	81.4	215	AAB30799	A modifie
6	485	36.0	721	ABP73437	Candida a
7	483	35.8	715	AAB30820	Amino aci
8	374.5	27.8	741	AAB30819	Amino aci
9	186	13.8	1145	AAB09770	Drosophil
10	186	13.8	1145	ADB38846	Drosophil
11	182.5	13.5	882	ABU19350	Protein e
12	175.5	13.0	407	ADP94879	ORF incor
13	169	12.5	720	ADB70240	C. neofor
14	168.5	12.5	405	AAB30810	Amino aci
15	168.5	12.5	405	AAB30821	Amino aci
16	167	12.4	592	ABO53050	Human put
17	167	12.4	592	AD126209	Human pro
18	163.5	12.1	589	AD126207	Human pro
19	153	11.4	952	ABR52717	Protein s
20	153	11.4	952	ADK61942	Disease t
21	151	11.2	492	ADJ69335	Human hea
22	151	11.2	1844	AAB18250	Plaemodiu
23	150.5	11.2	914	ABR52593	Protein s
24	150.5	11.2	914	ADK61944	Disease t
25	150	11.1	461	AAR75506	Nicotiana

26	149.5	11.1	1790	4	ABB71739	Drosophil
27	149.5	11.1	1833	8	AQO89584	Antagonis
28	148.5	11.0	207	5	ABP63118	FLOII gen
29	148	11.0	417	3	AAG30013	Arabidops
30	148	11.0	419	3	AAG30012	Arabidops
31	148	11.0	439	3	AAG30011	Arabidops
32	146.5	10.9	462	5	ABG95083	Human tra
33	146.5	10.9	462	8	ABE77159	Human pro
34	145	10.8	712	6	ABJ25450	Aspergill
35	145	10.8	716	6	ABJ26050	Aspergill
36	145	10.8	1162	3	AAV96255	Kaposi's
37	145	10.8	1162	3	AAV96255	Kaposi's
38	145	10.8	1162	4	AAV58500	HHV8 ORF
39	145	10.8	1162	4	ABE62331	Amino aci
40	145	10.8	1162	5	ABE05621	Kaposi's
41	144.5	10.7	738	8	ADJ65096	HHV8 late
42	143.5	10.6	391	8	ADN03304	Antipsori
43	143.5	10.6	392	4	ABG10550	Novel hum
44	143.5	10.6	404	2	AAR90675	Human chr
45	143	10.6	456	4	ABB65707	Drosophil

ALIGNMENTS

RESULT 1

AAB30792

ID AAB30792 standard; protein; 685 AA.

AC AAB30792;

XX 02-APR-2001 (first entry)

DE Amino acid sequence of a yeast Sup35 protein.

KW Sup35; phenotype: SCHAG; self-coalesce; higher-order aggregate;

KW amyloidogenic domain; aggregation; fibril; phenotypic alteration;

KW gene therapy; disease resistance; plant pigmentation; prion disease.

OS Saccharomyces cerevisiae.

PN WO200075324-A2.

PD 14-DEC-2000.

PF 09-JUN-2000; 2000WO-US015876.

PR 09-JUN-1999; 99US-0138833P.

XX (ARCH-) ARCH DEV CORP.

Lindquist S, Li L, Ma J, Liu J, Sondheimer N, Scheibel T;

WPI: 2001-061723/07.

DR N-PSDB; AAC86679.

New nucleic acid encoding chimeric proteins with self-assembly

properties, useful e.g. for diagnosis and treatment of prion diseases,

also related aggregates, fibrils and polymers.

XX Claim 11; Page 125-127; 189pp; English.

The present sequence represents a yeast Sup35 protein. The protein possesses the prion-like capacity to undergo a self-perpetuating conformational alteration that changes the functional state of Sup35 in a manner that creates a heritable change in phenotype. It is used to construct chimeric polypeptides of the invention, which comprise at least one SCHAG (self-coalesces into higher-order aggregates) amino acid sequence fused in frame with a polypeptide of interest (which is other than a marker protein, a glutathione-S-transferase or a staphylococcal nuclear protein). The specification also describes chimeric polypeptides that comprises an amyloidogenic domain that causes aggregation into fibrils. The chimeric polypeptides are used to prepare polymers with

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	269	20.0	159	4	US-09-248-796A-19824	Sequence 19824, A
2	230.5	17.1	139	4	US-09-248-796A-19825	Sequence 19825, A
3	151.5	11.2	1196	4	US-09-248-796A-19876	Sequence 19876, A
4	146.5	11.0	207	4	US-10-029-180-70	Sequence 70, Appl
5	148.5	10.9	462	4	US-09-919-039-324	Sequence 324, Appl
6	145	10.8	1162	2	US-08-728-223A-2	Sequence 2, Appl
7	145	10.8	1162	3	US-09-298-568-2	Sequence 2, Appl
8	145	10.8	1162	4	US-09-410-399-2	Sequence 2, Appl
9	145	10.8	1162	4	US-09-894-273-2	Sequence 2, Appl
10	143.5	10.6	391	4	US-09-949-016-6646	Sequence 6646, Ap
11	143.5	10.6	392	4	US-09-949-016-10150	Sequence 10150, A
12	142.5	10.6	1010	4	US-09-248-796A-16379	Sequence 16379, A
13	140.5	10.4	657	4	US-09-248-796A-19232	Sequence 19232, A
14	136	10.1	526	4	US-09-538-092-1080	Sequence 1080, Ap
15	135.5	10.1	545	4	US-09-248-796A-15777	Sequence 15777, A
16	134.5	10.0	210	4	US-09-248-796A-25835	Sequence 25835, A
17	134.5	10.0	588	4	US-09-248-796A-20839	Sequence 20839, A
18	134.5	10.0	980	4	US-09-248-796A-19242	Sequence 19242, A
19	134	9.9	284	4	US-09-502-540-12642	Sequence 12642, A
20	133	9.9	177	4	US-09-248-796A-19826	Sequence 19826, A
21	132.5	9.8	1020	4	US-09-538-092-911	Sequence 911, App
22	131	9.7	1179	4	US-09-949-016-7088	Sequence 7088, Ap
23	129.5	9.6	907	4	US-09-949-016-9750	Sequence 9750, Ap
24	129.5	9.6	907	4	US-09-949-016-9751	Sequence 9751, Ap
25	129	9.6	253	4	US-09-248-796A-19227	Sequence 19227, A
26	128	9.5	480	4	US-09-248-796A-23564	Sequence 23564, A
27	128	9.5	455	4	US-09-248-796A-20471	Sequence 20471, A

```

; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19825
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-19825

```

```

Query Match      17.1%; Score 230.5; DB 4; Length 139;
Best Local Similarity 40.7%; Pred. No. 1.1e-12;
Matches 66; Conservative 10; Mismatches 31; Indels 55; Gaps 7;

        63 YNPDAGYQ-QQYVPGGGVQQVNPQGGVQQQGFNPQGGRGYKYNFNNNLQGYQAGFQPS 121
           : . | | | | | | | | | | | | | | | | | | | | | | | | | | | |
        4 FQPAGEYNQGY-----GGYQQYNQGFQOQQO-----QQQS 33

       122 QGMSLNDPQKQ--QKQAA---PKPKTKLKVSSGGIKLANATKKVGT--KP----- 165
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
       34 QGMSLADFQKQKTEQOASLNKPAVKTKLQAGSSGIKLANATKKVDTTTSKFSQKSSPAP 93

       166 -----ARSDKKEEKSAETKEPTKPTVEEPVKKEKP 199
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
       94 APAPAAFPASQEKKEEKAATAATPETKNETSAPAEKTKETTP 135

```

RESULT 3
 US-09-248-796A-19876
 ; Sequence 19876, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 19876
 ; LENGTH: 1196
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (491)
 ; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
 US-09-248-796A-19876

```

Query Match      11.2%; Score 151.5; DB 4; Length 1196;
Best Local Similarity 26.5%; Pred. No. 0.00013;
Matches 62; Conservative 30; Mismatches 105; Indels 37; Gaps 10;

QY 11 QNYQOQYSONGNOQGNRYQGYQAYNAQAQPAQGYQYQNYQYSGYQ--QGYQOQYNPDAG 68
Db 4 KSPDDFKQKQKQSKPKFGFYRNN-----QGGYNTN-QAYGNVPTFGNQOQFOYQGG 57

QY 69 YQOQYNPQGYQOQYNPQGGYQOQFPNPGQGRGNKYKNFNNNLQGYQAGFOPQSGMSLN- 127
Db 58 YNQGFMQ--QYQYQGGYNNQGYNSNNRYNSNNYNNNTN-----QSGTISTPV 107

QY 128 DFOQKQQAAPKPKTKLKVSSGIGLANATKKVGTKPAESDKKEEKSAETKEPTKEPT 187
Db 108 DSLPTSGRSTPNAS--TTSLSLSLTALAKLVNSNIPFENLSNIEPKAKIAIRPE-- 163

QY 188 KVEEPKKEKPVQTEKTEKSELPKVEDLKISSESTHNTNNANVTSDALIKE 241
Db 164 -VETIVK-----IIDQEDLSIINBWKLNELKSLK---PKSPALVKE 203

```

```

RESULT 4
US-10-029-180-70
; Sequence 70, Application US/10029180
; Patent No. 6806082
; GENERAL INFORMATION:
; APPLICANT: Cali, Brian M.
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin T.
; APPLICANT: Milna, G. Todd
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeffry C.
; APPLICANT: Trueheart, Josh
; APPLICANT: Zhang, Lixin
; TITLE OF INVENTION: No. 6806082el Regulators of Fungal Gene Expression
; FILE REFERENCE: MIC-004
; CURRENT APPLICATION NUMBER: US/10/029,180
; CURRENT FILING DATE: 2001-12-22
; PRIOR APPLICATION NUMBER: US 60/257,431
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fungal gene
US-10-029-180-70

```

	Query Match	11.0%; Score 148.5; DB 4;	Length 207;	
	Best Local Similarity	36.5%;	Pred. No. 2.5e-05;	
	Matches	46; Conservative	8; Mismatches	47; Indels
			25; Gaps	8;
Qy	13 YQYSGNGNQCGNNRYGQTAVNAQPAGGY--YONYGYSGYGQGYYQY-NPDAGY	69		
Dd	4 YDOTNMGYGGQYGO-QGY-----GQPGYPGYGGQPYGQGHDOQQOYGPQHGY	56		
Qy	70 -QQYNPQGSYQY----NPQGGYQQFPNQGSRGNY-----KNFENYNNLQGYQAQF	117		
Dd	57 GGQYGGQGSDDYAGQHQHQQYGGQ---QGSSDYAGQHQQHGYSQHHDNRQGGY	113		
Qy	118 QPSQG	123		
Dd	114 EQOHG	119		

```

RESULT 5
US-09-919-039-324
; Sequence 324, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 324
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 1813444CD1
US-09-919-039-324

Query Match          10.9%; Score 146.5; DB 4; Length 462;
Best Local Similarity 24.0%; Pred. No. 0.0001;

```



```

; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Poly
US-09-538-092-1080

```

Query Match	10.1%;	Score 136;	DB 4;	Length 526;
Best Local Similarity	27.0%;	Pred. No. 0.001;		
Matches 51; Conservative	9;	Mismatches 57;	Indels 72;	Gaps 10;
QY	4	SNQNNQNTQQTSQNGNQO-QGNRRYQGY-----QAYNAQAQP-----AGGY-----	45	
Db	29	SSAPYGGQSYGYSQSTDTSYGGSSYSYSGQSQNTGYGTSTPQYGVSTGGYSSQSSQ	88	
QY	46	--YQNTQGYSGY-----QQGGYQYQYNPDAGYQQQYNPQGG	78	
Db	89	SSYGGQSSYPGYGQQAPFSTSGYSSSSSSSYGQPSGSYSQQPSYGGQQSYGQQ--	146	
QY	79	YQYNPQGGY--QQQFNPQ-----GGRGNY-----KNFNYYNNLQ--GYQ	114	
Db	147	QSYNPQGYGQQYQYNSSSGGGGGGGGGNYGQDQSSMSGSGSGGYGNQDQSGGGG	205	
QY	115	AGFQPSQ	123	
Db	206	GGYCGQDRG	214	

RESULT 15
 US-09-248-796A-15777
 ; Sequence 15777, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstein et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 15777
 ; LENGTH: 545
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 ; US-09-248-796A-15777

```

Query Match      10.1%; Score 135.5; DB 4; Length 545;
Best Local Similarity 26.0%; Pred. No. 0.0012;
Matches 56; Conservative 25; Mismatches 83; Indels 51; Gaps 11;

QY      3  DSNQGNQQNYQYQSYQNGNQQGNRRYQGYQAYNAQAQAGGYQNYQYSGYQGYQGYQ 62
DB      261  DLNDETTPQQPYLATGFYQQQPFY-----AQ-QPQQQPFQYDMF-----GNPI 306

QY      63  YNP-DAGYQQQYNPQGGYQYQNPQGGYQQQFNPQGGRGNYKNFYNNNNLQYQAGFQPS 121
DB      307  QNPMDTGL-----YNQAYYQCCQQQ-QQQQQFQP-----NQTFGNYG-----344

QY      122  QGMSLNDFOKQQAAPKPKKTKLKVSSSGIKLANATKKVGTAPAES--DKKBEKSAET 179
DB      345  -----QPQQQAQPEFPIQKTKSGNPFPMSSGSDNTNKPPTQSLSLAQEQQQQQQ 396

QY      180  KST--KEPTKVBEPVKBEKPVQTEKEEGEL 212
DB      397  QQPFQTQPTTA--PIKQNTSSSRNFETHELNDL 429

```

Search completed: July 1, 2005, 18:39:42
Job time : 20.5994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 18:20:02 ; Search time 66.5287 Seconds
(without alignments)
1467.128 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_253

Perfect score: 1348

Sequence: 1 MSDSNOGNQNYQQYSONG.....SADALIKEBEVDDEVND 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500.5	37.1	662	17	US-10-732-923-17478
2	485	36.0	721	14	US-10-032-585-7274
3	471.5	35.0	712	17	US-10-732-923-17486
4	413	30.6	701	17	US-10-732-923-17479
5	374.5	27.8	741	17	US-10-732-923-17107
6	300	22.3	742	17	US-10-732-923-17477
7	201.5	14.9	642	16	US-10-437-963-151492
8	186	13.8	1145	9	US-09-866-557A-5
9	186	13.8	1145	11	US-09-858-862-5
10	186	13.8	1145	14	US-10-055-797-5
11	186	13.8	1145	15	US-10-350-798-5

12	182.5	13.5	882	15	US-10-282-122A-47274	Sequence 47274, A
13	169	12.5	720	15	US-10-320-797-3284	Sequence 3284, Ap
14	155.5	11.5	2957	17	US-10-732-923-8692	Sequence 8692, Ap
15	153.5	11.4	671	16	US-10-437-963-109378	Sequence 109378, A
16	152	11.3	605	14	US-10-156-761-14380	Sequence 14380, A
17	151	11.2	492	16	US-10-408-765A-1141	Sequence 1141, Ap
18	150.5	11.2	914	16	US-10-477-369-63	Sequence 63, Appl
19	148.5	11.0	207	13	US-10-029-180-70	Sequence 70, Appl
20	148.5	11.0	207	17	US-10-952-045-70	Sequence 70, Appl
21	148	11.0	769	16	US-10-437-963-181417	Sequence 181417, A
22	146.5	10.9	462	10	US-09-919-039-324	Sequence 324, Ap
23	146.5	10.9	665	14	US-10-100-252-19	Sequence 19, Appl
24	145	10.8	712	14	US-10-128-714-3108	Sequence 3108, Ap
25	145	10.8	716	14	US-10-128-714-8108	Sequence 8108, Ap
26	145	10.8	1162	11	US-09-894-273-2	Sequence 2, Appl
27	145	10.8	1162	14	US-10-294-804-2	Sequence 2, Appl
28	145	10.8	1162	16	US-10-194-046-2	Sequence 238, App
29	144.5	10.7	738	16	US-10-451-467A-238	Sequence 184233, A
30	144	10.7	158	16	US-10-437-963-184233	Sequence 10929, A
31	143	10.6	320	16	US-10-739-930-10929	Sequence 7908, Ap
32	139	10.3	469	16	US-10-739-930-7908	Sequence 5663, Ap
33	139	10.3	632	16	US-10-739-930-5663	Sequence 650, App
34	138.5	10.3	525	16	US-10-755-889-550	Sequence 220613, A
35	137	10.2	440	15	US-10-424-599-220613	Sequence 496, App
36	135.5	10.1	292	16	US-10-451-467A-496	Sequence 7222, Ap
37	135.5	10.1	522	17	US-10-741-849-7222	Sequence 409, App
38	135	10.0	128	15	US-10-310-154-409	Sequence 514, App
39	135	10.0	128	17	US-10-732-923-514	Sequence 198246, A
40	135	10.0	412	16	US-10-425-115-198246	Sequence 148604, A
41	135	10.0	1474	16	US-10-437-963-148604	Sequence 148607, A
42	135	10.0	1514	16	US-10-437-963-148607	Sequence 3337, Ap
43	134	9.9	1197	17	US-10-732-923-3337	Sequence 211292, A
44	133.5	9.9	359	16	US-10-425-115-211292	Sequence 211329, A
45	133.5	9.9	477	16	US-10-425-115-211329	

ALIGNMENTS

RESULT 1

US-10-732-923-17478
; Sequence 17478, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 17478
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Zygosaccharomyces rouxii
US-10-732-923-17478

Query Match 37.1%; Score 500.5; DB 17; Length 662;
Best Local Similarity 51.5%; Pred. No. 1.4e-27;
Matches 134; Conservative 17; Mismatches 72; Indels 37; Gaps 15;

QY	1	MSDSNOGNQNYQQYSONGNOQGNRRYQGY-QAYNAQAQAPAGGYQNYQYSGYQOQG	59
Db	1	MSDPNQ-NGQ-----CGGQAGNRYQYFQKLTQQAQ-AGGYQPYGGY--GG	49
QY	60	YQYQNPADAGYQQYQNPQGGYQ-QYNPQGGYQQQFNPQGRGNKYNFNYYNNLQGYQAGFQ	118
Db	50	YGYQPYGGYQGYF--QDQQAQAGYNGYPYQ---AQAPGCGFN--NYYNQFQP-----Q	98
QY	119	POSQMSLNDFOKQO---KQAAPKPKYTKLVSSSGIKLANATKKVGTGTPAESDKKEBK	175
Db	99	QOSQMTLDDFHQKQTSQSPKPKKSLKLVSSSGIKLANATK-----KPKDEKKEBP	154


```

; Sequence 17107, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 17107
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Pichia pinus
US-10-732-923-17107

Query Match      27.8%; Score 374.5; DB 17; Length 741;
Best Local Similarity 37.4%; Pred. No. 1.6e-18;
Matches 105; Conservative 28; Mismatches 85; Indels 63; Gaps 14;

Qy      11 QNYQYQYSONGNOQGNRYQGYQAYNAQAQAGGYQNYQGYGYQGGYQOYQYNPDAGYQ 70
Db      56 QEQEQFGYQGGQQYNAQ--GGYNNYNNR-----GGYSNNRGYNNNRGGYSYNN---SYN 107
Qy      71 QYNYPQGGYQOYNPQGGYQQGFNPQGGRGNYKFNFNNN-----LQGYQAGF-----QPOSQ 122
Db      108 TNSN--QGGYSYNN-----NTYANNSYNNNNNNNNNNYNNQYNNYNSQPQGG 151
Qy      123 G-----MSLNDFOKQKQAA-----PKPKTKLKL--VSSSGIKLANATKKVGTKPAES 168
Db      152 DQOQETGSGQMSLEDYQKQKESLNKLTKEPKVKLNLSNSTVKAPIVTKKKEEPVQ 211
Qy      169 DKKEEKSAAE---TKPTKEPTKVEE-----PVKKEKPVQTEETKEKSG--EL 212
Db      212 ESKTEEPAKEEIKNOQEPAAENKVEESKVEAPTAAPVSESSEFPAST--PKTEAKASKEV 270
Qy      213 PKVEDLKISSTHTNNANVTISADALIKEQBEEVDDEVND 253
Db      271 AAAAAALKKEVSKQKESNVTNADALVKEQEQIDASTVND 311

```

```

RESULT 6
US-10-732-923-17477
; Sequence 17477, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 17477
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Yarrowia lipolytica
US-10-732-923-17477

Query Match          22.3%; Score 300; DB 17; Length 742;
Best Local Similarity 30.5%; Pred. No. 3.5e-13;
Matches 105; Conservative 24; Mismatches 91; Indels 124; Gaps 17;

Qy      1 MSDS-NQG--NNQQNYQQYSQ-NGNQOQQGNRRYQGYQAYNAAQ----- 40
        |||   |||   |||   |||   |||   |||   |||   |||   |||
Db      1 MSDFQNQDFANKANINEGQGGQGGQGG---QQGYGQAQPAQFVPGGSFVPFGQ 56
        |||   |||   |||   |||   |||   |||   |||   |||   |||

Qy      41 ---PAGGYQNYQGYSGVQ-----QGGYQQYNP-----DAGYQQQYNP 75
        |||   |||   |||   |||   |||   |||   |||   |||   |||
Db      57 SFA PQGGYDQ-YQGCGGYQGGYQGGYQGGYQGGYQGGYRGGRGGYRGGYNNQYGN 115
        |||   |||   |||   |||   |||   |||   |||   |||   |||

```

Qy 76 OGGYQOYNPOGGYQOQFNPOGG--PGNYKNFNYYNNNLQGYQAGFQPOSQMSLNDQKQOK 133
Db 116 OGGYQO--GYQGGYDNQYGNQGGYQGGY-----DNQYGYDQPEEPQEKTLTLEEYQKQKQ 168
Qy 135 QAAPKPKK-----TLKLVSSSGIKLANATKKVGTKPAES----- 168
Db 169 EALNKLKKPKVGGALKIGGDKPKAGGALKIGGDKRPASTGGALKIGGDKPKTSGGALKIGG 228
Qy 169 -----DKKEEKSAAETKEPT---KEPTKVEE---PVKKEKPVQTESEKTEBK 209
Db 229 DKPKSGGVLKIGGDKKESEKKEAVSPAPKESPTPEKSTPAPASKESTPSAAEKK--EA 286
Qy 210 SELPKVEDLKISESTHNTNNANVT SADALIKEQEEVEVDVND 253
Db 287 AE-----NNAASADALIAAQNABIDOEITKE 311

RESULT 7
US-10-437-963-151492
; Sequence 151492, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associa
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 151492
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_51630C.1.pep
US-10-437-963-151492

Query Match 14.9%; Score 201.5; DB 16; Length 642;
Best Local Similarity 31.5%; Pred. No. 3.3e-06;
Matches 70; Conservative 28; Mismatches 89; Indels 35; Gaps 10

Qy 7 GNNQONTQOYSQNGNQOQGNRRYQGYQYNAQAQAGGYQ--NYQGYSGYQGGYQOYNP 65
Db 438 GNNNGGYQQRGYNGN--NNGRFOQERAVNS---GNGGYRQGGYRGDGGYQQGRGYSNG 491
Qy 66 DAGYQ--QOYNPOGGYQO---YNPQGGYQOQFNPOGGRGNVKNFN-----YNNN 109
Db 492 NGGYQOQGGYSGNGGYQOQGRYSNNGGYRR---GNGOQNGRNESSDLSLPHFALTGT 547
Qy 110 LQGYQAGFQPOSQMSLNDQKQQAAPKPKTKLKLVSSSGIKLANA--TKKVGTKPAES 168
Db 548 IPATPEKVQSQSQASSSPAPQAQQA-----QTLSSSSSSGAAAPAPAKQAQSPAPPA 602
Qy 169 DKKEEKSAAETKEPTKTEPTKVEEPVKKEEKPQVTEKTEKS 210
Db 603 QAKSQAPPAQAK--SQAPAQVKSQAKAPAAARTKAPSKAQA 642

RESULT 8
US-09-866-557A-5
; Sequence 5, Application US/09866557A
; Patent No. US20020162126A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, S.
; APPLICANT: Hannon, G.
; APPLICANT: Beach, D.


```
Qy 6 GGNNOQ-----NYQY-----SONGNOGNNRYQGYQAYNAQAQAGGYQNYQY 52
Db 165 QGHGQGRGQGGYQQRPSGQGGHGHGQGRGQGGYQQRPPGQGGHGHGQGRGQGGG 224
Qy 53 -----SGYQGGYQYQ--YNPDAGYQOQYN--PQGGYQYQ--YNPQGGYQYQ--FNPQGGRG 99
Db 225 YQQRPSGQGGHGHGQGRGQGGYQQRPSGQGGHGHGQGRGQGGYQQRPSGQGGG-- 282
Qy 100 NYKNFNNNNLOGYQAGFQPOSGMSLNDPQK-----QQXQAAPKP 140
Db 283 -----HQGRGQGGYQQRPPGQPNQTSQGOYQSRGPPQQQQAAPLP 327

RESULT 12
US-10-282-122A-47274
; Sequence 47274, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47274
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-282-122A-47274

Query Match 13 5%; Score 182.5; DB 15; Length 882;
Best Local Similarity 27.0%; Pred. No. 0.00011;
Matches 80; Conservative 41; Mismatches 112; Indels 63; Gaps 15;

Qy 2 SDSNOG-----NNQYQYQYQY-----GNQOQGNRRYQGYQAYNAQAQAGGYQNYQY 52
Db 38 NDSNSFVLDHNNNSKAEYSQSRDNETGGYSQNRDNRAGGYS--QNRDNRAGGYSQNRDNR 95
Qy 53 S-GYQQ-----GGYQYQ--NPDAGYQOQYNPQGGYQYQ--YNPQGGY-QQOFPNPGG--- 97
```

```
Db 96 TGGYSQNRDNETGGYSQNRDNETGGYSQNRDNRGGYSQGRDNRRTGGYSQSRDNRRTGGYSQ 155
Qy 98 -----RGYKKNFNNNNLOGYQAGFQPOSGMSLNDPQKQKQKQAAAPKPKTK----- 145
Db 156 NRDNRTGGYSQ--NRDNRTGGYSQNRDNRRTGGYSQNRDLSLSPQYQGSVKTYVAKNSQNK 214
Qy 146 -----LVSSSGIKLANATKVKCTKPAESDKKEE--KSAETKEPTKE 185
Db 215 YTTTSMSPRRLIKTKVPAIVSSTPAADSENSKELNRKLGEKKKQOQESQSKYKTKKAETE 274
Qy 186 PTKVEEPVKKEEKPVQTEETEEKSELPKVDLKAISETHN--TNNANVTSSADALIK 240
Db 275 SKTIEQKVFEQ---LQKKKRENLANPIKSIDIMSGSIITVSDLARKMNLKSSDLIAK 327

RESULT 13
US-10-320-797-3284
; Sequence 3284, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Broshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3284
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
US-10-320-797-3284

Query Match 12.5%; Score 169; DB 15; Length 720;
Best Local Similarity 26.0%; Pred. No. 0.0008;
Matches 64; Conservative 22; Mismatches 80; Indels 80; Gaps 11;

Qy 38 QAQPAGGYQNYQGYSGYQGGYQOQYQYNPDPDAGYQOQYQYNPQGGYQOQYNPQGGYQOQY 97
Db 48 QQQPFPDPY-----GQQQGGYPQYQ-----QYQOQGGYPQYQYQGGYPQYQ----- 86
Qy 98 RGNYNFNNNNLOGYQA-----GFPQSQSGMSLNDPQKQKQKQAAAPKPKTKLTVS 148
Db 87 -----QQYVPVGPAGPAGPRAYQP--PQARNVQGFQPPSFSSPAPPPTKAPA 132
Qy 149 SSGIKLA-----NATKKVGTKPAESDKKEEKSAAETKEPTKVEEPVKKEE---KPVQ 201
Db 133 GKPVSLSISGGGAPK---AAPSLSTIEKEASSKSPKPAAPTTPKPAAPAAKSEASAPVS 189
Qy 202 TEETEEKSELPKVEDLK-----ISESTHNTNNANVT-----AD 236
Db 190 AAEKAKEKA-VPLTSDAQGKVVAETSAAKSGASTPVATVSTSTTTFKSVSAKNDAB 248
Qy 237 ALIKEQ 242
Db 249 AIYREQ 254

RESULT 14
US-10-923-923-8692
; Sequence 8692, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 18:06:15 ; Search time 14.7842 Seconds
(without alignments)
1646.548 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_253
Perfect score: 1348
Sequence: 1 MSDSNQGNQNYQYQSQNG.....SADALIKEBEVDVVD 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1348	100.0	685	1 EFBYS2	suppressor 2 prote
2	374.5	27.8	741	2 S12921	suppressor 2 prote
3	182.5	13.5	882	2 H70199	translation initia
4	171	12.7	729	2 T51896	probable translati
5	170.5	12.6	456	2 F8493	protein F57B9.9 [i
6	168.5	12.5	405	2 S19355	hypothetical prote
7	168.5	12.5	798	2 T33022	hypothetical prote
8	166	12.3	372	2 I64223	bifunctional endo-
9	163.5	12.1	589	2 S71954	RNA/ssDNA-binding
10	163.5	12.1	1094	2 S49313	protein kinase - s
11	161.5	12.0	3498	2 T22330	hypothetical prote
12	159.5	11.8	1655	2 T13998	gene mastermind pr
13	156	11.6	232	2 T21735	hypothetical prote
14	155.5	11.5	2957	2 T33152	hypothetical prote
15	153.5	11.4	409	2 T20847	hypothetical prote
16	153.5	11.4	1125	2 E90598	membrane nuclease,
17	153	11.4	462	1 LUD07	annexin VII - slim
18	153	11.4	952	2 S64473	translation initia
19	153	11.3	940	2 D89723	protein F39D8.1b [
20	152	11.3	945	2 T21998	hypothetical prote
21	151	11.2	1844	2 D71612	hypothetical prote
22	150.5	11.2	914	2 B48086	translation initia
23	150.5	11.2	1161	2 T18400	glutamate-cysteine
24	150.5	11.2	1390	2 T14004	trfA protein - eli
25	150	11.1	461	2 T10265	arabinogalactan-pr
26	149	11.1	765	2 T15447	hypothetical prote
27	148.5	11.0	867	2 T27136	hypothetical prote
28	148.5	11.0	871	2 T27135	hypothetical prote
29	148.5	11.0	1156	2 T23748	hypothetical prote

30	148	11.0	419	2	T04886	DAG protein homolo
31	147.5	10.9	352	2	T24389	hypothetical prote
32	147	10.9	597	2	T51889	related to ciathri
33	146.5	10.9	462	4	S33798	FUS/CHOP mutant fu
34	146.5	10.9	665	2	S70706	probable protein k
35	146.5	10.9	800	2	T51653	dsRNA-binding prot
36	145	10.8	1111	2	T29070	hypothetical prote
37	144.5	10.7	738	2	S37876	glutamine-rich pro
38	144.5	10.7	1004	2	H88562	protein C07A9.3 [i
39	144.5	10.7	1044	2	S40704	hypothetical prote
40	143.5	10.6	404	2	S46269	synovial sarcoma t
41	143.5	10.6	1500	2	T03824	probable immediate
42	142.5	10.6	284	2	T23158	hypothetical prote
43	142	10.5	676	2	S41022	hypothetical prote
44	140.5	10.4	236	2	T18240	hypothetical trans
45	140.5	10.4	870	2	T30110	hypothetical prote

ALIGNMENTS

RESULT 1

EFBYS2

suppressor 2 protein - yeast (Saccharomyces cerevisiae)

N;Alternate names: Gl-to-S transition protein; protein YD9395.05; protein YDR172w

C;Species: Saccharomyces cerevisiae

C;Date: 31-Dec-1991 Sequence revision 31-Dec-1991 #text change 09-Jul-2004

C;Accession: S00733; JT0323; S49768; S00488; A26742; S00533; S05723

R;Wilson, P.G.; Culbertson, M.R.

J. Mol. Biol. 199, 559-573, 1988

A;Title: SUP12 suppressor protein of yeast. A fusion protein related to the EF-1 family

A;Reference number: S00733; MUID:8817503; PMID:3280807

A;Accession: S00733

A;Molecule type: DNA

A;Residues: 1-685 <WIL>

A;Cross-references: UNIPROT:P05453; EMBL:X07163; NID:G4581; PIDN:CAA30155.1; PID:G4582

R;Kushnirov, V.V.; Ter-Avanesyan, M.D.; Telchov, M.V.; Surguchov, A.P.; Smirnov, V.N.; I

Gene 66, 45-54, 1988

A;Title: Nucleotide sequence of the SUP2(SUP35) gene of Saccharomyces cerevisiae.

A;Reference number: JT0323; MUID:88329727; PMID:3047009

A;Accession: JT0323

A;Molecule type: DNA

A;Residues: 1-685 <KUS>

A;Cross-references: EMBL:M21129; NID:G172789; PIDN:AAA35133.1; PID:G172791

R;Murphy, L.; Harris, D.E.

submitted to the EMBL Data Library, November 1994

A;Reference number: S49764

A;Accession: S49768

A;Molecule type: DNA

A;Residues: 1-685 <MUR>

A;Cross-references: EMBL:Z46727; NID:G1289283; PIDN:CAA86677.1; PID:G1289287; GSPDB:GN00

R;Kikuchi, Y.; Shimatake, H.; Kikuchi, A.

EMBO J. 7, 1175-1182, 1988

A;Title: A yeast gene required for the Gl-to-S transition encodes a protein containing a

A;Reference number: S00488; MUID:88296422; PMID:2841115

A;Accession: S00488

A;Molecule type: DNA

A;Residues: 1-52, 'C', 54-685 <KIK>

A;Cross-references: GB:Y00829; EMBL:Y00859; NID:G3711; PIDN:CAA68760.1; PID:G3712

C;Genetics: SGD:SUP35; SUP12; GST1; SUP2; MIPS:YDR172w

A;Gene: SGD:references: SGD:S0002579; MIPS:YDR172w

A;Map position: 4R

C;Superfamily: suppressor 2 protein; translation elongation factor Tu homology

C;Keywords: duplication; GTP binding; nucleotide binding; P-loop; phosphoprotein; tandem

F;1-123/Domain: A <DOM1>

F;42-119/Region: 10-residue repeats

F;124-253/Domain: charged <DOM2>

F;159-222/Region: glutamic acid/lysine-rich

F;254-685/Domain: C <DOM4>

F;261-409/Domain: translation elongation factor Tu homology <ETU>

F;267-274/Region: nucleotide-binding motif A (P-loop)

F;406-409/Region: GTP-binding NXKD motif

F:273/Binding site: GTP (Lys) #status predicted

```
Query Match      100.0%; Score 1348; DB 1; Length 685;
Best Local Similarity 100.0%; Pred. No. 4.2e-71;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSNGQNNQYQYQYQNGNQNNRYQGYQAYNAQAQAGGYQYQYQYQYQYQY 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MDSNGQNNQYQYQYQNGNQNNRYQGYQAYNAQAQAGGYQYQYQYQYQYQY 60

QY 61 QYNPDPAGYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 QYNPDPAGYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 120

QY 121 SQGMSLNDFOQKQQAAPKPKTKLKVSSSGIKLANATKVGTPKPAESDKKEEKSATK 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 SQGMSLNDFOQKQQAAPKPKTKLKVSSSGIKLANATKVGTPKPAESDKKEEKSATK 180

QY 181 EPTKEPTKVEBPVKKEEPVQTEETKESLKPVEDLKISESTHNTNNAVTSADALIK 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 EPTKEPTKVEBPVKKEEPVQTEETKESLKPVEDLKISESTHNTNNAVTSADALIK 240

QY 241 EQEEVDDVND 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 EQEEVDDVND 253

RESULT 2
S12921
N:Suppressor 2 protein - yeast (Pichia pinus)
A:Alternate names: SUP2 protein
C:Species: Pichia pinus
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 09-Jul-2004
C:Accession: S12921
R:Kushnir, V.V.; Ter-Avanesyan, M.D.; Didichenko, S.A.; Smirnov, V.N.; Chernoff, Y.O.;
A:Title: Divergence and conservation of SUP2 (SUP35) gene of yeasts Pichia pinus and Sacc
A:Reference number: S12921; MUID:91181341; PMID:2080663
A:Residues: 1-741 <KUS>
A:Molecule type: DNA
A:Cross-references: UNIPROT:P23637; EMBL:X56910; NID:g3235; PIDN:CAA40231.1; PID:g3236
C:Genetics:
A:Gene: SUP2
C:Superfamily: suppressor 2 protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop; phosphoprotein
F:319-467/Domain: translation elongation factor Tu homology <ETU>
F:325-332/Region: nucleotide-binding motif A (P-loop)
F:464-467/Region: GTP-binding NKXD motif
F:331/Binding site: GTP (Lys) #status predicted

Query Match      27.8%; Score 374.5; DB 2; Length 741;
Best Local Similarity 37.4%; Pred. No. 1.2e-14;
Matches 105; Conservative 28; Mismatches 85; Indels 63; Gaps 14;

QY 11 QNYQYQNGNQNNRYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 70
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 QOEQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 70

QY 71 QYNPQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 TNSN-QGYGNYN-----NNYANNYNNNNNNNNNNNNNNNNNNNNNNNNN 122

QY 123 G-----MSLNDFOQKQQA-----PKPKTKLKVSSSGIKLANATKVGTPKPAES 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 DQQQETSGGQMSLEDYQKQKESLKLNTKPKVYLKLNLSSTVKAPITVKKEEPEVQ 211

QY 169 DKKEEKSAS---TKETKEPTKVEE-----PVKKEEPVQTEETKES--EL 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 EKTEEPKAEIKNQEPAEAKNVEESKVEATPAKPVSESEFPAST-PKTEAKSKEV 270

QY 213 PKVEDLKISESTHNTNNAVTSADALIKQEVEEVDVND 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 271 AAAAAALKKEYSQAKKESNTVNADALYKEQEIQIDASIVND 311

RESULT 3

H70199

translation initiation factor IF-2 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C:Accession: H70199

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whitson, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: H70199

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-882 <KLE>

A:Cross-references: UNIPROT:O51741; GB:AE001179; GB:AE000783; NID:g2688738; PIDN:AAC671

A:Experimental source: strain B31

C:Keywords: GTP binding; nucleotide binding; P-loop

F:383-492/Domain: translation elongation factor Tu homology <ETU>

F:389-396/Region: nucleotide-binding motif A (P-loop)

F:489-492/Region: GTP-binding NKXD motif

F:525-527/Region: GTP-binding SAK/L motif

F:395,396,416,489,490,492,525/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser)

Query Match

Best Local Similarity 13.5%; Score 182.5; DB 2; Length 882;

Matches 80; Conservative 41; Mismatches 112; Indels 63; Gaps 15;

QY 2 SDSNQG---NNQY 52

Db 38 NDSNSFVDLHNSNKAESQSDNRNRTGGYQYQYQYQYQYQYQYQYQYQYQYQY 95

QY 53 S-GYQQ-----GGYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 97

Db 96 TGGYQY 155

QY 98 -----RGYKYNFNNNLNQYQAGFQYQYQYQYQYQYQYQYQYQYQYQYQY 145

Db 156 NRDNRTGGYQY 214

QY 146 -----LVSSSGIKLANATKVGTPKPAESDKKEE--KSAETKEPTKE 185

Db 215 YTTTSMFRLIKTKVPAIVSTPAADSENKELNKLGEKKQYQYQYQYQYQYQY 274

QY 186 PTKVEEPVKKEKPVQTEETKESLKPVEDLKISESTHNTNNAVTSADALIK 240

Db 275 SKTIEQKVFQ---LQKKRENLANPIPKSIDIMGSITVSDIARKNLKSLLIAK 327

RESULT 4

T51896

probable translation release factor erf3 [imported] - Neurospora crassa

N:Alternate names: protein B23111.80

C:Species: Neurospora crassa

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000

C:Accession: T51896

R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura

A:Reference number: Z25858

A:Accession: T51896

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-729 <SCH>

A:Cross-references: EMBL:AL391572; GSPDB:GN00116; NCSP:B23111.80

A:Experimental source: BAC clone B23111, strain OR74A

C:Genetics:

A:Gene: NCSP:B23111.80

A:Map position: 6

A:Introns: 111/1; 711/3
C:Superfamily: suppressor 2 protein; translation elongation factor Tu homology

Query Match 12.7%; Score 171; DB 2; Length 729;
Best Local Similarity 28.5%; Pred. No. 0.0073;
Matches 85; Conservative 23; Mismatches 130; Indels 60; Gaps 13;

QY 7 GNNQNYQ-----QYSONGNOQNNRYQGYQYNA----- 37
DB 3 GNVQNNWEEAADODERLARQTOQNNINAGTFR-PCAAATFGAPSFPTGQPAAPGFTPQ 61

QY 38 -QAQAGGYQNY-QGYSGYQGGYQOYNPDAGYQOQYQOYNPDAGYQOQYQOQYQOQYQO 95
DB 62 YQOQYVGAQYGGYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQO 119

QY 96 GGRGNY-KNFNYYNNLQYQAGFQSQGMSLND---FQKQQAAPK-----PKKTL---- 144
DB 120 QYQGYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQO 179

QY 145 -----KLVSSEGIKANATKVV---GTPAESDKKEEKSATKEPTKEPTKVPEPVKE 196
DB 180 GDAPAKVLSIGDAPKPAKVLISIGTAPA---KEPKKEAAKKEGTAEAAAKVATKAV 236

QY 197 EKPVOTEETEEKSELKPVEDLKISESTHTNNANVTSDALIKE-OEEVDEVVND 253
DB 237 QK----TESAASGRTPSPSSGRSPSAKSGKVSIRDVDVDEKDIQSADVEDTLKE 291

RESULT 5
F88493
protein F57B9.9 (imported) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F88493
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: AF5000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.banger.ac.uk/Projects/C.ele
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: F88493
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-456 <STO>
A:Cross-references: UNIPROT:Q20936; GB:chr_III; PIDN:AAA21172.1; PID:G532820; GSPDB:GN00
A:Gene: F57B9.9
A:Map position: 3

Query Match 12.6%; Score 170.5; DB 2; Length 456;
Best Local Similarity 27.0%; Pred. No. 0.0048;
Matches 67; Conservative 30; Mismatches 102; Indels 49; Gaps 13;

QY 13 YQYSONGNOQNNRYQGYQYNAQAQAPAGYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQ 72
DB 97 YAFVPOQYQOQ--NNFOQYGYPOQOQOQYQYNNNNNNNGY-----YNOQ---QOQ 144

QY 73 YNPOQGY-QOYNPQGYQOQY 131
DB 145 FN---GYNNYNNNNNNQOQY 193

QY 132 QOQQAAPKPKTLKLVSSSGIKLANATKVGTPAESDKKEEKSATKEPTKEPTKVPE 191
DB 194 QVQOQVQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQOQYQO 248

QY 192 PVKKEKP-----VOTEKTEKSELKPVEDLKISESTHTNNANVTSDALIKEQ 242
DB 249 LVQVIRPAASAVATAVVDTSKAKTEFNKANGIEQPKPAH-----AEKASLDA----- 299

QY 243 BEEVDEV 250
DB 300 --EVDANV 305

RESULT 6

S19355
hypothetical protein YCL028w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YCL181
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: S19355; S17471; C27061
R:Hollenberg, C.P.; Kleinhaus, U.; Lutzenkirchen, K.; Ramezani Rad, M.; Xu, G.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19350
A:Accession: S19355
A:Molecule type: DNA
A:Residues: 1-405 <HOI>
A:Cross-references: UNIPROT:P25367; EMBL:X59720; NID:G1907116; PIDN:CAA42357.1; PID:e864
R:Ramezani Rad, M.; Lutzenkirchen, K.; Xu, G.; Kleinhaus, U.; Hollenberg, C.P.
Yeast 7, 533-538, 1991
A:Title: The complete sequence of a 11,953 bp fragment from C1G on chromosome III encon
A:Reference number: S17471; MUID:91377317; PMID:1897318
A:Accession: S17471
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-186 <RAM>
A:Cross-references: EMBL:S58084
R:Trueheart, J.; Boeke, J.D.; Fink, G.R.
Mol. Cell. Biol. 7, 2316-2328, 1987
A:Title: Two genes required for cell fusion during yeast conjugation: evidence for a ph
A:Reference number: A93093; MUID:87286864; PMID:3302672
A:Accession: C27061
A:Molecule type: DNA
A:Residues: 1-180, 'T', 182-322, 'VHSQWLPLPQVPTWAIPIQIRVGCASNLMMSVDEN', 'RMVNSNPMSTEDNTA'
A:Cross-references: EMBL:M16717; NID:G171534; PIDN:AAA34615.1; PID:G171536
C:Genetics:
A:Gene: SGD.RNQ1
A:Cross-references: SGD:S0000533
A:Map position: 3L

Query Match 12.5%; Score 168.5; DB 2; Length 405;
Best Local Similarity 33.3%; Pred. No. 0.0055;
Matches 57; Conservative 10; Mismatches 49; Indels 55; Gaps 10;

QY 2 SDSNQNNQOYQOYSONGNOQNNRYQGYQYNAQAQAPAGYQYQYQYQYQYQYQYQYQY 54
DB 219 NNQNNNSOQYQ 271

QY 55 -----YQOQGYQOYNPDAGYQOQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQ 96
DB 272 SMAQSYLGGGTQSN-----QOQYQOQNNQOQYQOQYQOQYQOQYQOQYQOQYQO 323

QY 97 -----GRGNYKNFNYYNNLQYQAGFQSQG-MSLNDFOKQ 133
DB 324 SPSALASMASSYLGNNSNSYSGGQOQANEYG-RPOHNGQQQSNRYGRPQ 373

RESULT 7

T33022
hypothetical protein K07H8.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33022
R:Fulton, B.; Hawkins, J.; Gattung, S.; Wohldmann, P.; Elliott, G.
submitted to the EMBL Data Library, February 1998
A:Description: The sequence of C. elegans cosmid K07H8.
A:Reference number: Z21264
A:Accession: T33022
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-798 <FUL>
A:Cross-references: UNIPROT:Q45181; EMBL:AF047659; PIDN:AAC04430.1; GSPDB:GN000022; CESP
A:Experimental source: strain Bristol N2; clone K07H8
C:Genetics:
A:Gene: CESP.K07H8.10
A:Map position: 4

T33152
hypothetical protein T04D1.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33152
R:Davidson, S.; Wohldmann, P.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid T04D1.
A:Reference number: Z21292
A:Accession: T33152
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2957 <DAV>
A:Cross-references: UNIPROT:O61845; EMBL:AF067617; PIDN:AAC17559.1; GSPDB:GN00019; CESP:
A:Experimental source: strain Bristol N2; clone T04D1
C:Genetics:
A:Gene: CESP:T04D1.4
A:Map position: 1
A:Introns: 122/3; 293/3; 515/3; 1205/2; 1577/1; 2221/3; 2776/1; 2864/3
Query Match 11.5%; Score 155.5; DB 2; Length 2957;
Best Local Similarity 26.3%; Pred. No. 0.25;
Matches 70; Conservative 32; Mismatches 109; Indels 55; Gaps 12;
Qy 14 QQYS--QNGN-----QQGNNRYQGYQAYNAQAQAPAGGYQNYQYSGYQQGGYQ---- 61
Db 476 QQYNGMENGSGNYPVAQQQQQHPPMYAQYQQQQQ-----HAHQNNMMSGYPQQQQQPQHP 529
Qy 62 -QYNPDAGYQQQYNNPQGGYQYNNPQGGYQQQFPNPGQGRGNKYNANNLNQGYQAGFPQ- 119
Db 530 HQHQOPASVPMQHHPQHPHQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHP 580
Qy 120 -----SQSGHSLNDFQKQQAAPK-PKTKLKLVSSTSGGIKLANATKKVGTGKPA 166
Db 581 QSVDDQASPPQAQIPATST 636
Qy 167 ESKKKEEKAETKE--PTKEPTKVEEPVKKEKPVQTEE-KTEKSELKPKVEDLK---- 219
Db 637 VSDPIMPQSSNDQELQAIQSDQILEDPKPEEPIIEDEPKVEYSQETIPVEELKDEVE 696
Qy 220 ISESTHTNN--ANVTSADALIKEQE 243
Db 697 FKEDQQTSTNLEDDTGDIDNFVRNEE 722
RESULT 15
T20847
hypothetical protein F13E9.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20847
R:McMurray, A.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19332
A:Accession: T20847
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-409 <WIL>
A:Cross-references: UNIPROT:Q19414; EMBL:Z69383; PIDN:CAA93412.1; GSPDB:GN00022; CESP:FI
A:Experimental source: clone F13E9
C:Genetics:
A:Gene: CESP:F13E9.4
A:Map position: 4
A:Introns: 32/1; 275/3; 337/3
C:Superfamily: loricrin
Query Match 11.4%; Score 153.5; DB 2; Length 409;
Best Local Similarity 23.1%; Pred. No. 0.041;
Matches 72; Conservative 38; Mismatches 117; Indels 85; Gaps 13;
Qy 4 SNQGNQNTQQ-YSONGNCQQGNNRYQGYQAYNAQAQAPAGGYQYQNYQYSGYQ 56
Db 55 TGGGLLESSYQQGYGQNGSGMGYSQ-----QGYGNSQDDYGYISQSGSGMGVQGYGGSS 110

Qy 57 QGGYQYNNPDAGYQQQYNNPQGGYQ-----QYNPQGGYQQQYNNPQGGYQ 102
Db 111 QSYQQQ-----AFAQQQRPQQGFGSNGFSGQASGMNSFGQGGYQGNQNGFGGSGGFS 165
Qy 103 -----NFNYYNNLQGYQAGFQ-----PQSQG-----MSLNDQKQQ- 133
Db 166 QSGWGSNSLSSANSNGNN-QSSSSGYQNNQGRHQGGHSSSSSSSVMSNNGYSSNSG 224
Qy 134 --KQAAPKPKTKLVSSSGI-----KLANATKKVGTGKPAESDKKEE 173
Db 225 YGNNGGPTP-SFLNNVSSSAAQDYINIVNKSLLTTNQINEQASNWASANSVQAQYIQYET 283
Qy 174 EKSATKEPTKEPTKVEEPVKKEKPVQTEETEEKSELKPKVEDLKISETHNTNNAVT 233
Db 284 NRSAIHTQATQNVSDIINNLTATVOASIEGIQSNTSSSENDROSAFQALQQAHPMEYSVLT 343
Qy 234 SADALIKEQE 245
Db 344 S-----LQNQHEQ 351

Search completed: July 1, 2005, 18:21:02
Job time : 15.7842 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 18:05:25 ; Search time 40.2988 Seconds
(without alignments)
3214.888 Million cell updates/sec

Title: US-09-591-632-2_COPY_1_253
Perfect score: 1348
Sequence: 1 MSDSNQGNQNYQQYSQNG.....SADALIKEEEVDVND 253

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1348	100.0	685 1	ERF2_YEAST P05453 saccharomyc
2	1348	100.0	685 2	Q9HGV1 saccharomyc RN
3	1341	99.5	685 2	Q8TFB8 saccharomyc RX
4	1339	99.3	251 2	Q95TM0 saccharomyc RX
5	1339	99.3	429 2	Q96UJ2 saccharomyc RA
6	1332	98.8	251 2	Q96TL8 saccharomyc RA
7	1332	98.8	429 2	Q96TJ4 saccharomyc RT
8	1332	98.8	429 2	Q96UJ9 saccharomyc RL
9	1328	98.5	429 2	Q96UJ5 saccharomyc RN
10	1327	98.4	251 2	Q96UJ1 saccharomyc RP
11	1322	98.1	429 2	Q96UJ8 saccharomyc RX
12	1314	97.5	429 2	Q96UJ6 saccharomyc RX
13	1312	97.3	251 2	Q96UJ7 saccharomyc RT
14	1312	97.3	251 2	Q96UJ3 saccharomyc RT
15	1312	97.0	435 2	Q6Q7I3 saccharomyc RL
16	1308	97.0	251 2	Q96UJ8 saccharomyc RN
17	1308	97.0	429 2	Q96TQ9 saccharomyc RP
18	1305	96.8	435 2	Q6Q7I6 saccharomyc RX
19	1303	96.7	251 2	Q96UJ4 saccharomyc RX
20	1296	96.1	435 2	Q6Q7I2 saccharomyc RA
21	1290	95.7	435 2	Q6Q7I4 saccharomyc RA
22	1285	95.3	242 2	Q8TFR4 saccharomyc RT
23	1281	95.0	435 2	Q6Q7I1 saccharomyc RL
24	1279	94.9	243 2	Q8TFR3 saccharomyc RN
25	1275	94.6	240 2	Q8TFR0 saccharomyc RP
26	1262	93.6	237 2	Q8TFQ8 saccharomyc RC
27	1258	93.3	236 2	Q8TFQ9 saccharomyc RX
28	1241	92.1	235 2	Q8TFR1 saccharomyc RA
29	1240	92.0	233 2	Q8TFQ9 saccharomyc RA
30	1209.5	89.7	666 2	Q8TFR7 saccharomyc RA
31	1209.5	89.7	666 2	Q8TFR8 saccharomyc RA

32	1187	88.1	434	2	Q96UJ3	Q96UJ3 saccharomyc
33	1172.5	87.0	227	2	Q8TFR5	Q8TFR5 saccharomyc
34	1120.5	83.1	216	2	Q8TFR2	Q8TFR2 saccharomyc
35	701	52.0	123	2	Q8TFR6	Q8TFR6 saccharomyc
36	700	51.9	224	2	Q07089	Q07089 saccharomyc
37	686	50.9	120	2	Q8TFR9	Q8TFR9 saccharomyc
38	539.5	40.0	688	2	Q6FVM2	Q6FVM2 candida gla
39	500.5	37.1	662	2	Q9HGI4	Q9HGI4 zygoscacchar
40	483	35.8	299	2	Q9UVK3	Q9UVK3 candida alb
41	483	35.8	715	1	ERF2_CANAL	013354 candida alb
42	475	35.2	274	2	Q9UVJ9	Q9UVJ9 zygoscacchar
43	471.5	35.0	712	2	Q9HGI7	Q9HGI7 candida mal
44	456	33.8	691	2	Q750T4	Q750T4 ashbya goss
45	452	33.5	428	2	Q9URS3	Q9URS3 kluyveromyc

ALIGNMENTS

RESULT 1
ERF2_YEAST
ID ERF2_YEAST STANDARD; PRT; 685 AA.
AC P05453; P05420;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Eukaryotic peptide chain release factor GTP-binding subunit (ERF2)
DE (Translation release factor 3) (ERF3) (ERF-3) (Omnipotent suppressor
DE protein 2) (GI to S phase transition protein 1).
GN Name=SUP35; Synonyms=GST1, PNM2, SAL3, SUP12, SUP2;
GN OrderedLocusNames=YDR172W; ORFNames=YD9395.05;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycosaceae;
OX NCBI TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88329727; PubMed=3047009; DOI=10.1016/0378-1119(88)90223-5;
RA Kushnirov V.V., Ter-Avanesyan M.D., Telckov M.V., Surguchov A.P.,
RA Smirnov V.N., Inge-Vechtomov S.G.;
RT "Nucleotide sequence of the SUP2 (SUP35) gene of Saccharomyces
RT cerevisiae.";
RL Gene 66:45-54(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87219095; PubMed=3556215; DOI=10.1016/0014-5793(87)80157-6;
RA Kushnirov V.V., Ter-Avanesyan M.D., Surguchov A.P., Smirnov V.N.,
RA Inge-Vechtomov S.G.;
RT "Localization of possible functional domains in sup2 gene product of
RT the yeast Saccharomyces cerevisiae.";
RL FEBS Lett. 215:257-260(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88172503; PubMed=3280807;
RA Wilson P.G., Culbertson M.R.;
RT "SUP12 suppressor protein of yeast. A fusion protein related to the
RT EF-1 family of elongation factors.";
RL J. Mol. Biol. 199:559-573(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88296422; PubMed=2841115;
RA Kukuchi Y., Shimatake H., Kikuchi A.;
RT "A yeast gene required for the G1-to-S transition encodes a protein
RT containing an A-kinase target site and GTPase domain.";
RL EMBO J. 7:1175-1182(1988).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RX MEDLINE=97313263; PubMed=9169867;
RA Jacq C., Alt-Moerbe J., Andre B., Arnold W., Ballesta J.P.G.,
RA Bagues M., Baron L., Becker A., Bateau N., Bloeker H., Blugeon C.,
RA Boskovic J., Brandt P., Brueckner M., Buitrago M.J., Coster F.,
RA Delaveau T., del Rey F., Dujon B., Eide L.G., Garcia-Cantalejo J.M.,

Query Match	100.0%; Score 1348; DB 2; Length 685;
Best Local Similarity	100.0%; Pred. No. 8.3e-66;
Matches 253; Conservative	0; Mismatches 0; Indels 0; Gaps 0
Qy	1 MSDSNGNQNNQYQQYQSQNGNQQGNRRYQGYYAQAQAAGGYYQNYQGYSGYQQGGY 60
Dd	1 MSDSNGNQNNQYQQYQSQNGNQQGNRRYQGYYAQAQAAGGYYQNYQGYSGYQQGGY 60
Qy	61 QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQOQFNPGGRGNKYNFNNNNLQYQAQGFQPQ 120
Dd	61 QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQOQFNPGGRGNKYNFNNNNLQYQAQGFQPQ 120
Qy	121 SQGSLNDFOQOQOAAPKPKTLLKYSSSGIKLANATKKVGTKPAESDKKEEKSATK 180
Dd	121 SQGSLNDFOQOQOAAPKPKTLLKYSSSGIKLANATKKVGTKPAESDKKEEKSATK 180
Qy	181 EPTKEPTKVPEPVKKEEPVOTEETEEKSELPKVEDLKISESTHTNNANVTSDALIK 240
Dd	181 EPTKEPTKVPEPVKKEEPVOTEETEEKSELPKVEDLKISESTHTNNANVTSDALIK 240
Qy	241 EQEEVDDEVVND 253
Dd	241 EQEEVDDEVVND 253
RESULT 3	
Q8TFB8	PRELIMINARY; PRT; 685 AA.
ID	Q8TFB8
AC	Q8TFB8;
DT	01-JUN-2002 (TREMBLrel. 21, Created)
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE	Prión protein.
DN	Name=SUP35;
OS	Saccharomyces cerevisiae (Baker's Yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
RX	NCHI_TaxID=4932;
ON	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=SC12, and SC16;
RX	MEDLINE=22773310; PubMed=12890024;
RA	Resende C.G., Outeiro T.F., Sands L., Lindquist S., Tuite M.F.;
RT	"Prion protein gene polymorphisms in saccharomyces cerevisiae.";
RL	Mol. Microbiol. 49:1005-1017(2003).
DR	EMBL; AY028645; AAK26176.1; -
DR	EMBL; AY028649; AAK26180.1; -
DR	GCD; S000002579; SUP35.
DR	GO; GO:0005525; F:GTP binding; IEA.
DR	GO; GO:0003747; F:translation release factor activity; IEA.
DR	GO; GO:0006412; P:protein biosynthesis; IEA.
DR	GO; GO:0006415; P:translational termination; IEA.
DR	InterPro; IPR004160; EFTU_Cterm.
DR	InterPro; IPR004161; EFTU_D2.
DR	InterPro; IPR009001; Elong_init_C.
DR	InterPro; IPR000795; ProtSyn_GTPbind.
DR	InterPro; IPR009000; Translat_factor.
DR	InterPro; IPR003285; Yeast_ERF.
DR	Pfam; PF00009; GTP_EFTU_1.
DR	Pfam; PF03144; GTP_EFTU_D2; 1.
DR	Pfam; PF03143; GTP_EFTU D3; 1.
DR	PRINTS; PR00315; ELONGATNFCT.
DR	PRINTS; PR01343; YEASTERF.
DR	PROSITE; PS00301; EFACITOR_GTP; UNKNOWN 1.
KW	GTP-binding; Prión; Protein biosynthesis.
SW	SEQUENCE 685 AA; 76609 MW; 43912A6DDBAF4E53 CRC64;
Query Match	99.5%; Score 1341; DB 2; Length 685;
Best Local Similarity	99.6%; Pred. No. 2e-65;
Matches 252; Conservative	0; Mismatches 1; Indels 0; Gaps 0
Qy	1 MSDSNGNQNNQYQQYQSQNGNQQGNRRYQGYYAQAQAAGGYYQNYQGYSGYQQGGY 60

Db	1	MSDSNQGNQNYQYQYQSGNQOQGNRRYQGYQAYNAQAQAGGYQNYQGYSGYQQGGY	60
Qy	61	QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNTKKNFYNNNNLQGYQAGFPQP	120
Db	61	QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNTKKNFYNNNNLQGYQAGFPQP	120
Qy	121	SOGNSLNDFOKQQAAPKPKTKLKYSSSGIKLANATKKVGTGPAESDKKEEKSATK	180
Db	121	SOGNSLNDFOKQQAAPKPKTKLKYSSSGIKLANATKKVGTGPAESDKKEEKSATK	180
Qy	181	EPTKEPTKVVEPVKKEKPVQTEETKSELSPLKVEDLKI SESTHNTNNANVTSADALIK	240
Db	181	EPTKEPTKVVEPVKKEKPVQTEETKSELSPLKVEDLKI SESTHNTNNANVTSADALIK	240
Qy	241	EQBEVDDEVVND 253	
Db	241	EQBEVDDEVVND 253	
RESULT 4			
Q96TMO	ID	Q96TMO PRELIMINARY; PRT; 251 AA.	
AC	Q96TMO;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	Sup35p (fragment).		
GN	Name=SUP35;		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
OX	NCBI_TaxID=4932;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=W303, 74D694, S288C, YJM145, and YJM210;		
RC	MEDLINE=21518585; PubMed=11606530;		
RA	Jensen M.A., True H.L., Chernoff Y.O., Lindquist S.;		
RT	"Molecular population genetics and evolution of a prion-like protein		
RL	in <i>Saccharomyces cerevisiae</i> ."		
RL	Genetics 159:527-535 (2001).		
DR	EMBL; AY052599; AAL15014.1; -		
DR	EMBL; AY052600; AAL15015.1; -		
DR	EMBL; AY052601; AAL15016.1; -		
DR	EMBL; AY052603; AAL15018.1; -		
DR	EMBL; AY052604; AAL15019.1; -		
FT	NON_TER 1		
FT	NON_TER 251		
SQ	SEQUENCE 251 AA; 28300 MW; 75FDE507AD427068 CRC64;		
Query Match 99.3%; Score 1339; DB 2; Length 251;			
Best Local Similarity 100.0%; Pred. No. 9.3e-66;			
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0			
Qy	3	DSNQGNQNYQYQYQSGNQOQGNRRYQGYQAYNAQAQAGGYQNYQGYSGYQQGGYQQ	62
Db	1	DSNQGNQNYQYQYQSGNQOQGNRRYQGYQAYNAQAQAGGYQNYQGYSGYQQGGYQQ	60
Qy	63	YNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNTKKNFYNNNNLQGYQAGFPQSQ	122
Db	61	YNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNTKKNFYNNNNLQGYQAGFPQSQ	120
Qy	123	GMSLNDFOKQQAAPKPKTKLKYSSSGIKLANATKKVGTGPAESDKKEEKSATK	182
Db	121	GMSLNDFOKQQAAPKPKTKLKYSSSGIKLANATKKVGTGPAESDKKEEKSATK	180
Qy	183	TKEPTKVVEPVKKEKPVQTEETKSELSPLKVEDLKI SESTHNTNNANVTSADALIKEQ	242
Db	181	TKEPTKVVEPVKKEKPVQTEETKSELSPLKVEDLKI SESTHNTNNANVTSADALIKEQ	240
Qy	243	EEVDDEVVND 253	
Db	241	EEVDDEVVND 251	

Qy 1 MSDSNQGNQNYQQYTSQNGNQOQNNRYQGYQAYNAQAPAGGYQNYQGYSGYQGGY 60

```

RESULT 5
Q96UJ2 ID Q96UJ2 PRELIMINARY; PRT; 429 AA.
AC Q96UJ2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sup35p (Fragment).
GN Name=Sup35;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=SL1010-1A;
RC MEDLINE=21518585; PubMed=11606530;
RA Jensen M.A., True H.L., Chernoff Y.O., Lindquist S.;
RT "Molecular population genetics and evolution of a prion-like protein
in Saccharomyces cerevisiae.";
RL Genetics 159:527-535(2001).
DR EMBL; AY052613; AAL15028.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0003747; F:translation release factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0008415; P:translational termination; IEA.
DR InterPro; IPR000795; Protsyn_GTPbind.
DR Pfam; PF00009; GTP_EFTU; 1.
DR PRINTS; PR00315; YEASTERP.
DR PROSITE; PS00301; EFACOR_GTP; UNKNOWN_1.
KW GTP-binding; Protein biosynthesis.
FT NON_TER 1
FT NON_TER 429
FT NON_TER 429
SQ SEQUENCE 429 AA; 48238 MW; BAEC0BC1D6B8A3B2 CRC64;

Query Match 99.3%; Score 1339; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.6e-65;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DSNQGNQNYQYQSQNGNQGNRYQGYQAYNAQAQAGGYQNYQGYSGYQGGYQQ 62
Db 1 DSNQGNQNYQYQSQNGNQGNRYQGYQAYNAQAQAGGYQNYQGYSGYQGGYQQ 60

Qy 63 YNPAGYQQYQNPQGGYQQYQNPQGGYQQQFNPQGGRGYKNFNYNLNQYQAGFQPSQ 122
Db 61 YNPAGYQQYQNPQGGYQQYQNPQGGYQQQFNPQGGRGYKNFNYNLNQYQAGFQPSQ 120

Qy 123 GMSLNDFOKQQAAPKPKTKLVSSSGIKLANATKKVGTKPASDKKEEKSATKEP 182
Db 121 GMSLNDFOKQQAAPKPKTKLVSSSGIKLANATKKVGTKPASDKKEEKSATKEP 180

Qy 183 TKEPTKVEEPVKKEKPVQTEETKESLKPVEDLKISESTHNTNNAVTSADALIKEQ 242
Db 181 TKEPTKVEEPVKKEKPVQTEETKESLKPVEDLKISESTHNTNNAVTSADALIKEQ 240

Qy 243 EEEVDDEVND 253
Db 241 EEEVDDEVND 251

RESULT 6
Q96TL8 ID Q96TL8 PRELIMINARY; PRT; 251 AA.
AC Q96TL8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Sup35p (Fragment).
GN Name=Sup35;
OS Saccharomyces cerevisiae (Baker's yeast).

```

```

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=YJM273, YJM434, and YJM454;
RC MEDLINE=21518585; PubMed=11606530;
RA Jensen M.A., True H.L., Chernoff Y.O., Lindquist S.;
RT "Molecular population genetics and evolution of a prion-like protein
in Saccharomyces cerevisiae.";
RL Genetics 159:527-535(2001).
DR EMBL; AY052605; AAL15020.1; -.
DR EMBL; AY052607; AAL15022.1; -.
DR EMBL; AY052610; AAL15025.1; -.
FT NON_TER 1
FT NON_TER 251
FT NON_TER 251
SQ SEQUENCE 251 AA; 28358 MW; 770235049E41405B CRC64;

Query Match 98.8%; Score 1332; DB 2; Length 251;
Best Local Similarity 99.6%; Pred. No. 2.2e-65;
Matches 250; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DSNQGNQNYQYQSQNGNQGNRYQGYQAYNAQAQAGGYQNYQGYSGYQGGYQQ 62
Db 1 DSNQGNQNYQYQSQNGNQGNRYQGYQAYNAQAQAGGYQNYQGYSGYQGGYQQ 60

Qy 63 YNPAGYQQYQNPQGGYQQYQNPQGGYQQQFNPQGGRGYKNFNYNLNQYQAGFQPSQ 122
Db 61 YNPAGYQQYQNPQGGYQQYQNPQGGYQQQFNPQGGRGYKNFNYNLNQYQAGFQPSQ 120

Qy 123 GMSLNDFOKQQAAPKPKTKLVSSSGIKLANATKKVGTKPASDKKEEKSATKEP 182
Db 121 GMSLNDFOKQQAAPKPKTKLVSSSGIKLANATKKVGTKPASDKKEEKSATKEP 180

Qy 183 TKEPTKVEEPVKKEKPVQTEETKESLKPVEDLKISESTHNTNNAVTSADALIKEQ 242
Db 181 TKEPTKVEEPVKKEKPVQTEETKESLKPVEDLKISESTHNTNNAVTSADALIKEQ 240

Qy 243 EEEVDDEVND 253
Db 241 EEEVDDEVND 251

RESULT 7
Q96TJ4 ID Q96TJ4 PRELIMINARY; PRT; 429 AA.
AC Q96TJ4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Sup35p (Fragment).
GN Name=Sup35;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=REDSTAP and YJM308;
RC MEDLINE=21518585; PubMed=11606530;
RA Jensen M.A., True H.L., Chernoff Y.O., Lindquist S.;
RT "Molecular population genetics and evolution of a prion-like protein
in Saccharomyces cerevisiae.";
RL Genetics 159:527-535(2001).
DR EMBL; AY052616; AAL15031.1; -.
DR EMBL; AY052619; AAL15034.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0003747; F:translation release factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0008415; P:translational termination; IEA.
DR InterPro; IPR000795; Protsyn_GTPbind.
DR Pfam; PF00009; GTP_EFTU; 1.

```


DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Translation termination factor SUP35 (fragment).
GN Name: SUP35;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M22, S288C, and M8;
RX PubMed=15059259;
RA Fay J.C., McCullough H.L., Sniegowski P.D., Eisen M.B.;
RT "Population genetic variation in gene expression is associated with
RT phenotypic variation in Saccharomyces cerevisiae.";
RL Genome Biol. 5:R26-R26(2004).
DR EMBL; AY553987; AAS64328.1; -;
DR EMBL; AY553990; AAS64331.1; -;
DR EMBL; AY553985; AAS64326.1; -;
DR GO; GO:000525; P:GTP binding; IEA.
DR GO; GO:0003747; P:translation release factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006415; P:translational termination; IEA.
DR InterPro; IPR00795; ProtSyn_GTPbind.
DR PRINTS; PR00315; ELONGATNFT.
DR PRINTS; PR01343; YEASTERF.
DR PROSITE; PS00301; EFACITOR_GTP; UNKNOWN_1.
KW GTP-binding; Protein biosynthesis.
FT NON_TER 1
FT NON_TER 435
SQ SEQUENCE 435 AA; 48968 MW; ED0E85699BB28442 CRC64;

Query Match 97.3%; Score 1312; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 4.8e-64;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 NNQNYQYQSQNGNQGNRYQGYAYNAQAQAGGYQNYQGYQGGYQGYQYNPDA 67
DB 1 NNQNYQYQSQNGNQGNRYQGYAYNAQAQAGGYQNYQGYQGGYQGYQYNPDA 60

QY 68 GYQQYNPQGGYQYQYNPQGGYQQQFNPQGGRGYKNFNNNLNQGYQAGFPQSQGMSLN 127
DB 61 GYQQYNPQGGYQYQYNPQGGYQQQFNPQGGRGYKNFNNNLNQGYQAGFPQSQGMSLN 120

QY 128 DFQKQQAAPKPKTKLVSSGKLANATKVTGTPAESDKKEEKSATKEPTKEPT 187
DB 121 DFQKQQAAPKPKTKLVSSGKLANATKVTGTPAESDKKEEKSATKEPTKEPT 180

QY 188 KVEEPVKKEKPVQTEETKESKELPKVEDLKISESTHTNNANVTSDALIKEQEEVD 247
DB 181 KVEEPVKKEKPVQTEETKESKELPKVEDLKISESTHTNNANVTSDALIKEQEEVD 240

QY 248 DEVVND 253
DB 241 DEVVND 246

Search completed: July 1, 2005, 18:19:53
Job time : 41.2988 secs

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)